

/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-483-577A-3-	7.00	96.79	1.1e+03	5009	!	/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-295-802-140+	6.00	113.34	126.23	40
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-837-438-3-	7.00	96.79	1.1e+03	5009	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-868-968C-7+	6.00	113.34	126.23	40
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-937-654-3-	7.00	96.79	1.1e+03	5009	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-488-337A-140+	6.00	113.34	126.23	40
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-487-890A-2-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-375-992A-140+	6.00	112.99	126.17	42
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-337-483-2-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-185-949B-98+	6.00	112.99	133.12	40
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-478-373-2-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-449-045C-25+	6.00	111.71	155.76	50
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-478-373-2-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-330-867A-41+	6.00	111.71	155.76	50
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-483-577A-2-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-348-260-41+	6.00	111.71	155.76	50
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-637-654-2-	7.00	96.76	1.1e+03	5033	!	/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-956-047-2-	6.00	110.74	176.23	57
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-390-721-1-	7.00	96.74	1.1e+03	5045	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-185-850-307+	6.00	109.67	202.32	66
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-487-890A-4-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-680-326-44-	6.00	108.44	236.81	78
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-478-435-4-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-172-711-61-	6.00	106.99	285.14	95
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-337-483-4-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-178-015-5+	6.00	106.99	285.14	95
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-483-577A-4-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-337-658A-48+	6.00	103.69	439.26	100
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-897-438-4-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-484-570A-48+	6.00	103.69	435.70	149
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-637-654-4-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-888-072A-38+	6.00	103.64	435.46	150
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-837-654A-4-	7.00	96.66	1.1e+03	5039	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-478-403-10+	6.00	103.54	443.96	152
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-478-435-105-	7.00	96.60	1.1e+03	5144	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-478-403-10+	6.00	103.54	443.96	152
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-483-577A-105-	7.00	96.60	1.1e+03	5144	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-337-678-18-	6.00	101.09	607.38	122
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-897-438-105-	7.00	96.60	1.1e+03	5144	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-457-304A-28-	6.00	101.09	607.38	122
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-637-654-105-	7.00	96.60	1.1e+03	5144	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-457-304A-28-	6.00	101.09	607.38	122
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-977-405A-1+	7.00	96.49	1.1e+03	5220	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-456-701A-18-	6.00	101.09	607.38	122
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-977-871A-1+	7.00	96.49	1.1e+03	5220	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-456-701A-18-	6.00	101.09	607.38	122
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-09-1225-951-1+	7.00	96.49	1.1e+03	5220	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-456-701A-36+	6.00	101.06	610.08	122
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-1225-951-1+	7.00	96.08	1.2e+03	5157	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-456-701A-36+	6.00	101.06	610.08	122
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-1225-951-1+	7.00	96.08	1.2e+03	5157	!	/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-456-701A-36+	6.00	101.06	610.08	122
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-450-332-1-	7.00	90.65	2.3e+03	11561	!	/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-934-877A-31+	6.00	101.06	610.08	122
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-637-640-1-	7.00	90.65	2.3e+03	11561	!	/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-934-877A-31+	6.00	101.06	610.08	122
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-637-640-1-	7.00	90.65	2.3e+03	11561	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-871-878C-41+	6.00	101.06	610.08	122
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-871-878C-41+	6.00	100.69	639.72	224
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-367-841A-43-	7.00	85.77	4.3e+03	22481	!	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08-700-575-19-	6.00	100.12	688.04	244
/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-08										

cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-702-344-11	6.00	97.67	942.54	338	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-473-031-14	6.00	97.52	960.92	345	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-473-031-14	6.00	97.52	960.92	345	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-060-756-61	6.00	97.29	989.76	356	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-136-277-21	6.00	96.52	1.1e+03	395	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-479-403-21	6.00	96.52	1.1e+03	395	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-835-734-21	6.00	96.52	1.1e+03	395	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-07-846-595-2	6.00	96.47	1.1e+03	398	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-469-555-2	6.00	96.47	1.1e+03	398	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-318-193-88	6.00	96.41	1.1e+03	401	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-318-193-88	6.00	96.21	1.1e+03	412	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-227-357-71	6.00	96.20	1.1e+03	413	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-196-388-1	6.00	96.02	1.2e+03	423	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-323-873A-11	6.00	95.77	1.2e+03	438	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-387-942C-38	6.00	95.60	1.2e+03	448	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-680-395-8	6.00	95.49	1.2e+03	453	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-680-395-8	6.00	95.49	1.2e+03	455	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-060-756-605	6.00	95.42	1.3e+03	459	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-184-009-139	6.00	95.37	1.3e+03	462	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-458-356-139	6.00	95.37	1.3e+03	462	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-09-012-366-1	6.00	95.37	1.3e+03	462	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-828-735-159	6.00	95.37	1.3e+03	462	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-160-567-10	6.00	95.25	1.3e+03	470	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-031-143B-73	6.00	95.25	1.3e+03	470	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-031-143B-73	6.00	94.94	1.3e+03	490	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-060-756-587	6.00	94.91	1.3e+03	492	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-5219739-21	6.00	94.82	1.4e+03	498	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-928-637-55	6.00	94.78	1.4e+03	501	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-09-004-838-154	6.00	94.75	1.4e+03	503	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-545-809A-21	6.00	94.52	1.4e+03	519	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-473-031-11	6.00	94.35	1.4e+03	531	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-348-572-6	6.00	93.84	1.5e+03	569	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-09-041-090B-6	6.00	93.84	1.5e+03	569	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-07-938-188-1	6.00	93.80	1.5e+03	572	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-306-062-1	6.00	93.80	1.5e+03	572	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-190-560-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5A.COMB	seq:US-08-469-277-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-469-277-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata/2/1lna/5B.COMB	seq:US-08-468-946-3	6.00	93.72	1.6e+03	579	1	5
cgcn2_6/p.toddata							

/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-08-925-091-6-	6.00	89.16	2.8e+03	1076	/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-09-177-776-49+	6.00	87.22	3.6e+03	1400
/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-09-077-675A-6+	6.00	89.16	2.8e+03	1076	/cgn2_6/p.toddata/2/ina/5a.COMB.seq:US-08-146-424-19+	6.00	87.21	3.6e+03	1400
/cgn2_6/p.toddata/2/ina/5a.COMB.seq:US-08-022-411-1-	6.00	89.06	2.8e+03	1088	/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-08-693-448-13+	6.00	87.21	3.6e+03	1404
/cgn2_6/p.toddata/2/ina/5a.COMB.seq:US-09-716-161A-3-	6.00	89.03	2.8e+03	1091	/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-08-419-709-31+	6.00	87.21	3.6e+03	1400
/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-08-944-604-19-	6.00	89.03	2.9e+03	1096	/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-09-233-613-60-	6.00	87.21	3.6e+03	1407
/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-09-716-161A-10-	6.00	88.92	2.9e+03	1112	/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-08-879-260-1-	6.00	87.19	3.6e+03	1400
/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-08-767-675A-9+	6.00	88.86	2.9e+03	1122	/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-08-605-430-69+	6.00	87.18	3.6e+03	1400
/cgn2_6/p.toddata/2/ina/5b.COMB.seq:US-08-480-203-1+	6.00	88.62	3.0e+03	1158	/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-366-704C-50+	6.00	86.94	3.7e+03	1400
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-09-027-357-41+	6.00	88.62	3.0e+03	1158	/cgn2_6/p.toddata/2/ina/5B.COMB.seq:US-08-923-613-60-	6.00	86.92	3.7e+03	1400
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-09-056-556-16+	6.00	88.54	3.0e+03	1171	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-976-215-3+	6.00	86.91	3.7e+03	1462
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-107-676-26-	6.00	88.39	3.1e+03	1178	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-09-776-221-3+	6.00	86.91	3.7e+03	1462
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-890-719-10-	6.00	88.39	3.1e+03	1195	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-004-838-126-	6.00	86.90	3.7e+03	1488
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-018-584A-37-	6.00	88.36	3.1e+03	1200	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-08-330-121B-1-	6.00	86.81	3.8e+03	1488
/cgn2_6/p.toddata/2/ina/5B.COMB.seq:US-08-107-676-2-	6.00	88.30	3.1e+03	1211	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-08-828-701A-3+	6.00	86.81	3.8e+03	1488
/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-08-229-287-3-	6.00	88.23	3.2e+03	1221	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-08-828-701A-3+	6.00	86.81	3.8e+03	1488
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-890-719-2+	6.00	88.18	3.2e+03	1230	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-09-160-567-3+	6.00	86.77	3.8e+03	1490
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-990-719-9+	6.00	88.18	3.2e+03	1230	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-09-160-567-3+	6.00	86.77	3.8e+03	1490
/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-08-240-372-1+	6.00	88.14	3.2e+03	1237	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-833-876-1-	6.00	86.77	3.8e+03	1491
/cgn2_6/p.toddata/2/ina/5B.COMB.seq:US-08-591-629-1+	6.00	88.04	3.2e+03	1253	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-632-470-51-	6.00	86.77	3.8e+03	1491
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-642-274D-49-	6.00	88.04	3.2e+03	1254	/cgn2_6/p.toddata/2/ina/5B.COMB.seq:US-08-349-035-1+	6.00	86.72	3.8e+03	1501
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-952-014C-49-	6.00	88.00	3.3e+03	1261	/cgn2_6/p.toddata/2/ina/5B.COMB.seq:US-08-566-650A-1+	6.00	86.72	3.8e+03	1501
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-775-742-16+	6.00	88.00	3.3e+03	1261	/cgn2_6/p.toddata/2/ina/5B.COMB.seq:US-08-566-650A-1+	6.00	86.72	3.8e+03	1501
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-989-386-4+	6.00	87.99	3.3e+03	1262	/cgn2_6/p.toddata/2/ina/5B.COMB.seq:US-09-010-546-23+	6.00	86.72	3.8e+03	1501
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-858-207A-162+	6.00	87.95	3.3e+03	1269	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-09-010-715-23+	6.00	86.72	3.8e+03	1501
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-378-58B-23-	6.00	87.95	3.3e+03	1269	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-577-780B-23+	6.00	86.72	3.8e+03	1501
/cgn2_6/p.toddata/2/ina/5B.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-944-916-1+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-817-877-1+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6B.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5A.COMB.seq:US-09-173-915-12+	6.00	86.56	3.9e+03	1533
/cgn2_6/p.toddata/2/ina/6A.COMB.seq:US-08-811-094-23-	6.00	87.95	3.3e+03	1270	/cgn2_6/p.toddata/2/ina/5				

[illegible]


```

/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-652-446-34- 6.00 83.75 5.6e+03 2249 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-910-820-7+ 6.00 83.74 5.6e+03 2251 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-322-742-2+ 6.00 83.71 5.6e+03 2261 1
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-09-276-400-3+ 6.00 83.71 5.6e+03 2262 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-448-076-3+ 6.00 83.70 5.7e+03 2263 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-176-126B-1- 6.00 83.70 5.7e+03 2263 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-669-435-1- 6.00 83.70 5.7e+03 2263 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:PCR-US94-14431A-1- 6.00 83.70 5.7e+03 2263 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-444-005-14+ 6.00 83.69 5.7e+03 2268 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-09-197-360-1+ 6.00 83.67 5.7e+03 2273 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-168-629-1+ 6.00 83.67 5.7e+03 2273 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-810-131A-1+ 6.00 83.60 5.7e+03 2295 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-842-842-6- 6.00 83.60 5.7e+03 2295 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-09-052-521C-1- 6.00 83.52 5.8e+03 2319 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-286-767-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-090-526-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-486-715-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-486-719-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-271-354-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-476-100-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-565-861-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-475-749-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:PCR-US94-07658-1- 6.00 83.52 5.8e+03 2320 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-08-975-762-44- 6.00 83.51 5.8e+03 2322 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-295-028-44- 6.00 83.51 5.8e+03 2322 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-106-582-44- 6.00 83.50 5.8e+03 2322 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-08-868-435-1+ 6.00 83.50 5.8e+03 2327 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-08-744-231-1+ 6.00 83.50 5.8e+03 2327 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-08-811-481-34- 6.00 83.49 5.8e+03 2338 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-557-128-3- 6.00 83.49 5.8e+03 2330 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-09-242-690A-34- 6.00 83.46 5.8e+03 2330 1
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-09-022-983-4+ 6.00 83.46 5.8e+03 2340 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-07-821-716-3- 6.00 83.41 5.9e+03 2356 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-07-923-724-7+ 6.00 83.39 5.9e+03 2363 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-609-426A-7+ 6.00 83.39 5.9e+03 2363 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-981-650-1- 6.00 83.38 5.9e+03 2369 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-883-534-4+ 6.00 83.37 5.9e+03 2369 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-09-204-764-4+ 6.00 83.37 5.9e+03 2369 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-374-652C-1+ 6.00 83.34 5.9e+03 2379 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-468-558-4+ 6.00 83.33 5.9e+03 2382 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-676-444-4- 6.00 83.33 5.9e+03 2382 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-319-989-7- 6.00 83.29 6.0e+03 2394 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-08-311-023-7+ 6.00 83.26 6.0e+03 2404 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-490-099-1- 6.00 83.19 6.0e+03 2427 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-09-624-693A-18+ 6.00 83.15 6.1e+03 2439 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-08-858-207A-45- 6.00 83.13 6.1e+03 2445 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-07-863-169A-4+ 6.00 83.08 6.1e+03 2464 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-429-964-4+ 6.00 83.08 6.1e+03 2464 1
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-07-935-087-4+ 6.00 83.08 6.1e+03 2464 1
/cgn2_6/pdata/2/ina/6A.COMB.seq:PCR-US93-08062-4+ 6.00 83.07 6.1e+03 2464 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-09-490-692-3+ 6.00 83.03 6.2e+03 2477 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-630-118A-1+ 6.00 83.03 6.2e+03 2481 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-838-399-1+ 6.00 83.03 6.2e+03 2481 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-235-893-1+ 6.00 83.02 6.2e+03 2485 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-483-043-3+ 6.00 82.99 6.2e+03 2492 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-481-238-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-471-066B-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-484-956-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/5A.COMB.seq:US-08-757-653-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-599-491-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-756-386-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/5B.COMB.seq:US-08-823-516-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-682-853A-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-759-038-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/6A.COMB.seq:US-08-758-316-3+ 6.00 82.96 6.2e+03 2504 1
/cgn2_6/pdata/2/ina/6B.COMB.seq:US-09-350-309-3+ 6.00 82.96 6.2e+03 2504 1

```


[illegible]

/cg2_6/pdata/2/ina/5B_COMB.seq:US-08-975-762-42 - 6.00 79.92 9.1e+03 3786
/cg2_6/pdata/2/ina/5B_COMB.seq:US-09-295-028-42 - 6.00 79.92 9.1e+03 3786
/cg2_6/pdata/2/ina/5B_COMB.seq:US-09-106-582-42 - 6.00 79.92 9.1e+03 3786
/cg2_6/pdata/2/ina/5A_COMB.seq:US-07-872-644-42 - 6.00 79.92 9.1e+03 3789
/cg2_6/pdata/2/ina/5A_COMB.seq:US-08-297-494-42 - 6.00 79.92 9.1e+03 3789
/cg2_6/pdata/2/ina/5A_COMB.seq:US-08-454-455-5 + 6.00 79.92 9.1e+03 3789
/cg2_6/pdata/2/ina/5A_COMB.seq:US-08-479-532-42 - 6.00 79.92 9.1e+03 3789
/cg2_6/pdata/2/ina/5A_COMB.seq:US-08-455-526-42 - 6.00 79.92 9.1e+03 3789
/cg2_6/pdata/2/ina/5A_COMB.seq:US-08-455-525-42 - 6.00 79.92 9.1e+03 3789
/cg2_6/pdata/2/ina/5A_COMB.seq:US-09-139-491-42 - 6.00 79.92 9.1e+03 3789
/cg2_6/pdata/2/ina/5A_COMB.seq:US-09-139-491-42 - 6.00 79.92 9.1e+03 3789
/cg2_6/pdata/2/ina/5B_COMB.seq:US-08-936-165A-196 - 6.00 79.92 9.1e+03 3815
/cg2_6/pdata/2/ina/5B_COMB.seq:US-08-823-120-5

seq_documentation_block:
Sequence 5, Application US/08823120
Patent No. 6149919
GENERAL INFORMATION:
APPLICANT: Domenighini, Mario
APPLICANT: Rappunli, Rino
APPLICANT: Pizze, Mariagrazia
TITLE OF INVENTION: Immunogenic Detoxified Mutants of
TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and
TITLE OF INVENTION: Their Use for the Preparation of Vaccines
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,120
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,003
FILING DATE: 11-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET INFORMATION: 0315,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
NAME/KEY: CDS
LOCATION: 1..711
US-08-823-120-5

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-823-120-5 ..

Align seg 1/1 to: US-08-823-120-5 from: 1 to: 711
103 ValTyrSerProHisProTyrGluGluGluValSerAlaLeuGlyIle 119
119 eProTyrSerGlnIleTyrGlyTyrTyrArgValAspPheGlyValIle 136
345 ACCATATTCACAGATATATGATGATGATGCTGTAAATTTCGTAGTGGAT 394
136 spGluArgLeuHisArgAsnArgGluTyrArgAspArgIleTyrArgAsn 152
395 ATGACACATTACATCCGTACACGGGAATATAGACACCGGTATTCAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhePro 169
445 CTGATATATACCTCCGCGACAGATGGTTACAGATATAGACGTTCCACC 494
169 CAspHisGlnAlaTArpArgGluGluProTyrIleHisAlaProGln 186
495 GATCACACCAAGCTTGGAGAGAACCCCTGATTTCATCATGCACCAAG 544
186 LysGlyGly 188
545 GTTGTGGA 552

seq_name: /cg2_6/pdata/2/ina/5B_COMB.seq:US-08-435-605A-7

seq_documentation_block:
Sequence 7, Application US/08435605A
Patent No. 5874287
GENERAL INFORMATION:
APPLICANT: Burnette, W. Neal
APPLICANT: Kaslow, Harvey R.
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
TITLE OF INVENTION: SUBUNIT ANALOGS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,605A
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196B
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-435-605A-7

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-435-605A-7 ..

Align seg 1/1 to: US-08-435-605A-7 from: 1 to: 576

29 GUUyrpheaSPARGlyThrGlnmetAsnIleasnLeuTyraSPHISAl 45
|||||
85 GAGTACTTGGACCGAGTACTCAATGATATCAACCTTATGATCATGC 134
|||||
45 aargGlyThrGlnThrglyPheValArg 54
|||||
135 AAGAGGAACCTCAGACGGGATTTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-435-605A-5

seq_documentation_block:

Sequence 5, Application US/08435605A
Patent No. 5874287
GENERAL INFORMATION:
APPLICANT: Burnette, W. Neal
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
SUBUNIT ANALOGS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,605A
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196B
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 582 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-435-605A-5

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-435-605A-5 ..

Align seg 1/1 to: US-08-435-605A-5 from: 1 to: 582

29 GUUyrpheaSPARGlyThrGlnmetAsnIleasnLeuTyraSPHISAl 45
|||||
85 GAGTACTTGGACCGAGTACTCAATGATATCAACCTTATGATCATGC 134
|||||
45 aargGlyThrGlnThrglyPheValArg 54
|||||
135 AAGAGGAACCTCAGACGGGATTTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-823-120-7

seq_documentation_block:

Sequence 7, Application US/08823120
Patent No. 6149919
GENERAL INFORMATION:
APPLICANT: Domenighini, Mario
APPLICANT: Rappuoli, Rino
APPLICANT: Pizza, Mariagrazia
TITLE OF INVENTION: Immunogenic Detoxified Mutants of
Cholera Toxin and of the Toxin B, Their Preparation and
USE FOR THE PREPARATION OF VACCINES
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,120
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,003
FILING DATE: 11-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0315,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..723
US-08-823-120-7

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-823-120-7 ..

Align seg 1/1 to: US-08-823-120-7 from: 1 to: 723

29 GUUyrpheaSPARGlyThrGlnmetAsnIleasnLeuTyraSPHISAl 45
|||||
85 GAGTACTTGGACCGAGTACTCAATGATATCAACCTTATGATCATGC 134
|||||
45 aargGlyThrGlnThrglyPheValArg 54
|||||
135 AAGAGGAACCTCAGACGGGATTTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-435-605A-1

seq_documentation_block:

Sequence 1, Application US/08435605A


```
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196B
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-435-605A-1

alignment_scores:
  Quality: 26.00      Length: 26
  Ratio: 1.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x US-08-435-605A-1 ..
  Align seg 1/1 to: US-08-435-605A-1 from: 1 to: 774

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
  |||||||
139 GAGTACTTTCACCGAGGTACTCAATGAATATCAACCTTATGATCATGTC 188
45 aArgGlyThrGlnThrGlyPheValArg 54
  |||||||
189 AAGAGGAACCTCAGACGGGATTGTGTTAGG 216

seq_name: /cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-449-045C-1

seq_documentation_block:
; Sequence 1, Application US/08449045C
; Patent No. 5770203
; GENERAL INFORMATION:
; APPLICANT: Burnette, Neal W.
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,045C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,733
; FILING DATE: 02-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,222
; FILING DATE: 06-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196C
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..777
; US-08-449-045C-1
```

```
alignment_scores:
  Quality: 26.00      Length: 26
  Ratio: 1.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x US-08-449-045C-1 ..
  Align seg 1/1 to: US-08-449-045C-1 from: 1 to: 777
```

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
  |||||||
139 GAGTACTTTCACCGAGGTACTCAATGAATATCAACCTTATGATCATGTC 188
45 aArgGlyThrGlnThrGlyPheValArg 54
  |||||||
189 AAGAGGAACCTCAGACGGGATTGTGTTAGG 216

seq_name: /cgn2_6/ptodata/2/ina/backfiles1.seq:5223610-1

seq_documentation_block:
; Patent No. 5223610
; APPLICANT: Burton, Frank H.; Sutcliffe, Gregor
; TITLE OF INVENTION: CHOLERA TOXIN GENE REGULATED BY GROWTH
; HORMONE PROMOTER
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/528,852
; FILING DATE: 18-MAY-1990
; SEQ ID NO: 1:
; LENGTH: 2020
; 5223610-1
```

```
alignment_scores:
  Quality: 26.00      Length: 26
  Ratio: 1.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x 5223610-1 ..
```


Align seg 1/1 to: 5223610-1 from: 1 to: 2020

29 Gltutyrheasparglythrghlmetasnleasnleutyraaphisai 45
|||||
654 GAGTACTTTGACCGAGTACTCAATGATATCACTTATGATCATGTC 703
45 aargglythrghlthrghlphvalarg 54
|||||
704 AAGAGCACTCAGACGGGATTTGTAGG 731

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-296-848A-1

seq_documentation_block:

Sequence 1, Application US/08296848A
Patent No. 601982

GENERAL INFORMATION:

APPLICANT: Clements, John D.

APPLICANT: Dickinson, Bonny L.

TITLE OF INVENTION: MUTANT ENTEROTOXIN EFFECTIVE AS A

TITLE OF INVENTION: NON-TOXIC ORAL ADJUVANT

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESSEE: PENNIE & EDMONDS LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/296,848A

FILING DATE: 26-AUG-1994

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Mistock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 5113-046

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA

US-08-296-848A-1

alignment_scores:

Quality: 15.00 Length: 15

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-296-848A-1 ..

Align seg 1/1 to: US-08-296-848A-1 from: 1 to: 45

186 GtYcsglyasnserargthrllthrghlvaspthrCysasn 200
|||||

1 GGTGTGGAATTCATCATCAGACAACTTACGTCATCTGTAAAT 45

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-9

seq_documentation_block:

Sequence 9, Application US/08435605A

Patent No. 5874287

GENERAL INFORMATION:

APPLICANT: Burnette, W. Neal

APPLICANT: Kaslow, Harvey R.

TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN

TITLE OF INVENTION: SUBUNIT ANALOGS

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Amgen Inc.

STREET: 1840 De Havilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/435,605A

FILING DATE: 05-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mazza, Richard J.

REGISTRATION NUMBER: 27,657

REFERENCE/DOCKET NUMBER: A-196B

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 138 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-435-605A-9

alignment_scores:

Quality: 12.00 Length: 12

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-435-605A-9 ..

Align seg 1/1 to: US-08-435-605A-9 from: 1 to: 138

213 GltutyrghnserlysrvalrglnrlpheSer 224
|||||

55 GAATACCAATCTAAAGTTAAAGACAAATATTTTCA 90

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-449-045C-23

seq_documentation_block:

Sequence 23, Application US/08449045C

Patent No. 5770203

GENERAL INFORMATION:

APPLICANT: Burnette, Neal W.

APPLICANT: Kaslow, Harvey R.

TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN

TITLE OF INVENTION: SUBUNIT ANALOGS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Amgen Inc.

STREET: 1840 De Havilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible


```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,045C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,733
; FILING DATE: 02-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,222
; FILING DATE: 06-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196C
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; 08-449-045C-23

alignment_scores:
      Quality: 11.00      Length: 11
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-449-045C-23 ..
Align seg 1/1 to: US-08-449-045C-23 from: 1 to: 50

179 TrpIleHisHisAlaProGingIyCysGlyAsn 189
|||||
3 TGGATTCATCATGACCGCAGCGGTGTGGGAAT 35

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-449-045C-24

seq_documentation_block:
; Sequence 24, Application US/08449045C
; Patent No. 5770203
; GENERAL INFORMATION:
; APPLICANT: Burnette, Neal W.
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,045C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/694,733
; FILING DATE: 02-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/271,222
; FILING DATE: 06-JUL-1994
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196C
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; 08-449-045C-24

alignment_scores:
      Quality: 11.00      Length: 11
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-449-045C-24 ..
Align seg 1/1 to: US-08-449-045C-24 from: 1 to: 50

179 TrpIleHisHisAlaProGingIyCysGlyAsn 189
|||||
3 TGGATTCATCATGACCGCAGCGGTGTGGGAAT 35

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-435-605A-35

seq_documentation_block:
; Sequence 35, Application US/08435605A
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-196B
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Oligonucleotide"
; 08-435-605A-35

alignment_scores:
      Quality: 11.00      Length: 11
      Ratio: 1.000      Gaps: 0
```


Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-435-605A-35 ..

Align seg 1/1 to: US-08-435-605A-35 from: 1 to: 50

179 TPTLEHSHSLAPROGLNGLYCysGLYasn 189
|||||
3 TGGATTCATCATGCCGCGAGGTTGTGGAAAT 35

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-435-605A-36

seq_documentation_block:

; Sequence 36, Application US/08435605A

; Patent No. 5874287

GENERAL INFORMATION:

APPLICANT: Burnette, W. Neal

APPLICANT: Kaslow, Harvey R.

TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN

TITLE OF INVENTION: SUBUNIT ANALOGS

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: 1840 De Havilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/435, 605A

FILING DATE: 05-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Mezza, Richard J.

REGISTRATION NUMBER: 27,657

REFERENCE/DOCKET NUMBER: A-196B

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "oligonucleotide"

US-08-435-605A-36

alignment_scores:

Quality: 11.00 Length: 11

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-435-605A-36 ..

Align seg 1/1 to: US-08-435-605A-36 from: 1 to: 50

179 TPTLEHSHSLAPROGLNGLYCysGLYasn 189
|||||
3 TGGATTCATCATGCCGCGAGGTTGTGGAAAT 35

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-629-600-14

seq_documentation_block:

; Sequence 14, Application US/08629600

; Patent No. 5783196

GENERAL INFORMATION:

APPLICANT: NORIEGA, Fernando

APPLICANT: LEVINE, Myron M.

TITLE OF INVENTION: GUA MUTANTS OF SHIGELLA

TITLE OF INVENTION: AND VACCINES CONTAINING THE SAME

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W., Suite 800

CITY: Washington, D.C.

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/629,600

FILING DATE: 9-APR-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: KIP, Gordon

REGISTRATION NUMBER: 30,764

REFERENCE/DOCKET NUMBER: A-6765

TELEPHONE: (202) 293-7060

TELEFAX: (202) 293-7860

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 42 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

HYPOTHETICAL: NO

US-08-629-600-14

alignment_scores:

Quality: 10.00 Length: 10

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-629-600-14 ..

Align seg 1/1 to: US-08-629-600-14 from: 1 to: 42

192 ARGTHRIETHRGVYASPTHCysASnGLU 201
|||||
13 AGAACAAATTACAGGATCTGTATATGAG 42

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-823-120-36

seq_documentation_block:

; Sequence 36, Application US/08823120

; Patent No. 6149919

GENERAL INFORMATION:

APPLICANT: Domenighini, Mario

APPLICANT: Rappuoli, Rino

APPLICANT: Pizzi, Mariagrazia

TITLE OF INVENTION: Immunogenic Detoxified Mutants of

TITLE OF INVENTION: Cholera toxin and of the toxin Lt, Their Preparation and

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: California

COUNTRY: USA

ZIP: 94608-2916


```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,120
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,003
FILING DATE: 11-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0315,001
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-823-120-36
```

```
alignment_scores:
Quality: 8.00      Length: 8
Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-08-823-120-36  ..
Align seg 1/1 to: US-08-823-120-36 from: 1 to: 28
```

```
8 Alaaspsarargproaspglu 15
|||||
2 GCAGATCTAGACCTCTGATGAA 25
```

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-403-584-5

seq_documentation_block:

```
; Sequence 5, Application US/08403584
; Patent No. 5631010
; GENERAL INFORMATION:
; APPLICANT: Mekalanos, John J.
; TITLE OF INVENTION: DELETION MUTANTS AS VACCINES
; TITLE OF INVENTION: FOR CHOLERA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,584
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,388
; FILING DATE:
; APPLICATION NUMBER: US/07/909,382
```

```
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00742/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 32
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-403-584-5
```

```
alignment_scores:
Quality: 8.00      Length: 8
Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-08-403-584-5  ..
Align seg 1/1 to: US-08-403-584-5 from: 1 to: 32
```

217 LysValLysArgGlnIlePheSer 224

```
|||||
6 AAAGTTAAAGACCAATATTTTCA 29
```

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-367-115-5

seq_documentation_block:

```
; Sequence 5, Application US/08367115
; Patent No. 5874088
; GENERAL INFORMATION:
; APPLICANT: Mekalanos, John J.
; APPLICANT: Beattie, David
; APPLICANT: Killen, Kevin
; APPLICANT: Lu, Yichen
; TITLE OF INVENTION: DELETION MUTANTS AS VACCINES FOR CHOLERA
; FILE REFERENCE: 00742/002002
; CURRENT APPLICATION NUMBER: US/08/367,115
; CURRENT FILING DATE: 1995-01-05
; EARLIER APPLICATION NUMBER: PCT/US93/06270
; EARLIER FILING DATE: 1993-07-01
; EARLIER APPLICATION NUMBER: 07/909,382
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Vibrio cholerae
US-08-367-115-5
```

```
alignment_scores:
Quality: 8.00      Length: 8
Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-08-367-115-5  ..
Align seg 1/1 to: US-08-367-115-5 from: 1 to: 32
```

217 LysValLysArgGlnIlePheSer 224

```
|||||
6 AAAGTTAAAGACCAATATTTTCA 29
```



```
seq_name: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq:PCT-US93-06270-5
;
; seq_documentation_block:
;   Sequence 5, Application PC/TUS9306270
;   GENERAL INFORMATION:
;     APPLICANT: Mekalanos, John J.
;     TITLE OF INVENTION: DELETION MUTANTS AS VACCINES FOR
;     TITLE OF INVENTION: CHOLERA
;     NUMBER OF SEQUENCES: 6
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Fish & Richardson
;       STREET: 225 Franklin Street
;       CITY: Boston
;       STATE: Massachusetts
;       COUNTRY: U.S.A.
;       ZIP: 02110-2804
;
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 3.5" Diskette, 1.44 MB
;     COMPUTER: IBM PS/2 Model 502 or 555X
;     OPERATING SYSTEM: MS-DOS (Version 5.0)
;     SOFTWARE: Wordperfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: PCT/US93/06270
;       FILING DATE: 19930701
;       CLASSIFICATION:
;       PRIORITY APPLICATION DATA:
;         APPLICATION NUMBER:
;         FILING DATE:
;         ATTORNEY/AGENT INFORMATION:
;           NAME: Freeman, John W.
;           REGISTRATION NUMBER: 29,066
;           REFERENCE/DOCKET NUMBER: 00742/007001
;           TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 542-5070
;             TELEFAX: (617) 542-8906
;             TELEX: 200154
;           INFORMATION FOR SEQ ID NO: 5:
;             SEQUENCE CHARACTERISTICS:
;               LENGTH: 32
;               TYPE: nucleic acid
;               STRANDEDNESS: single
;               TOPOLOGY: linear
;             PCT-US93-06270-5
;
;   alignment_scores:
;     Quality: 8.00      Length: 8
;     Ratio: 1.000      Gaps: 0
;     Percent Similarity: 100.000      Percent Identity: 100.000
;
;   alignment_block:
;     US-09-528-682-1 x PCT-US93-06270-5
;
;   Align seq 1/1 to: PCT-US93-06270-5 from: 1 to: 32
;
;   217 LysValLysArgGlnIlePheSer 224
;   ||||||||||||||||||||
;   6 AAAGTTAAAGACAATATTTTCA 29
;
; seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-296-848A-4
;
; seq_documentation_block:
;   Sequence 4, Application US/08296848A
;   Patent No. 6019982
;   GENERAL INFORMATION:
;     APPLICANT: Clements, John D.
;     APPLICANT: Dickinson, Bonny L.
;     TITLE OF INVENTION: MUTANT ENTEROTOXIN EFFECTIVE AS A
;     TITLE OF INVENTION: NON-TOXIC ORAL ADJUVANT
;     NUMBER OF SEQUENCES: 5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: PENNIE & EDMONDS LLP
;       STREET: 1155 Avenue of the Americas
```

```
;
;   CITY: New York
;   STATE: New York
;   COUNTRY: U.S.A.
;   ZIP: 10036
;
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/296,848A
;       FILING DATE: 26-AUG-1994
;       CLASSIFICATION: 424
;       ATTORNEY/AGENT INFORMATION:
;         NAME: Mistrock, S. Leslie
;         REGISTRATION NUMBER: 18,872
;         REFERENCE/DOCKET NUMBER: 5113-046
;         TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (212) 790-9090
;           TELEFAX: (212) 869-8864/9741
;           TELEX: 66141 PENNIE
;       INFORMATION FOR SEQ ID NO: 4:
;         SEQUENCE CHARACTERISTICS:
;           LENGTH: 45 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: double
;           TOPOLOGY: unknown
;         MOLECULE TYPE: DNA
;
;   US-08-296-848A-4
;
;   alignment_scores:
;     Quality: 8.00      Length: 8
;     Ratio: 1.000      Gaps: 0
;     Percent Similarity: 100.000      Percent Identity: 100.000
;
;   alignment_block:
;     US-09-528-682-1 x US-08-296-848A-4
;
;   Align seq 1/1 to: US-08-296-848A-4 from: 1 to: 45
;
;   193 ThrIleThrGlyAspThrCysAsn 200
;   ||||||||||||||||||||
;   22 ACAATTCAGGTGATCTGTAT 45
;
; seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-09-360-186-1
;
; seq_documentation_block:
;   Sequence 1, Application US/09360186
;   Patent No. 6262343
;   GENERAL INFORMATION:
;     APPLICANT: Staskawicz, et al.
;     TITLE OF INVENTION: Bst2 Resistance Gene
;     FILE REFERENCE: 50687
;     CURRENT APPLICATION NUMBER: US/09/360,186
;     CURRENT FILING DATE: 1999-07-23
;     EARLIER APPLICATION NUMBER: 60/093,957
;     EARLIER FILING DATE: 1998-07-23
;     NUMBER OF SEQ ID NOS: 9
;     SOFTWARE: PatentIn Ver. 2.0
;     SEQ ID NO 1
;     LENGTH: 31491
;     TYPE: DNA
;     ORGANISM: Capsicum annuum
;
;   US-09-360-186-1
;
;   alignment_scores:
;     Quality: 8.00      Length: 8
;     Ratio: 1.000      Gaps: 0
;     Percent Similarity: 100.000      Percent Identity: 100.000
;
;   alignment_block:
```


US-09-528-682-1 x US-09-360-186-1/rev ..
Align seg 1/1 to reverse of: US-09-360-186-1 from: 1 to: 31491

114 SerialaLeuGlyIleProTyr 121
|||||
6271 TCTGGCTAGGTGATACCCCTAC 6248

seq_name: /cgn2_6/ptodata/2/1na/5A.COMB.seq:US-07-932-379A-11

seq_documentation_block:
; Sequence 11, Application US/07932379A
; Patent No. 5468852
; GENERAL INFORMATION:
; APPLICANT: Ohashi, Tetsuo
; APPLICANT: Toda, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ozaki, Hiroko
; APPLICANT: Nishimura, Nasyuki
; APPLICANT: Shirasaki, Yoshinari
; APPLICANT: Yamagata, Koichi
; TITLE OF INVENTION: Oligonucleotides for Detecting
; TITLE OF INVENTION: Bacteria and Detection Method
; TITLE OF INVENTION: Using Same
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/932,379A
; FILING DATE: 19920819
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 1327-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli H10407
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..22
; OTHER INFORMATION: /label=oligonucleotide
; ;
; OTHER INFORMATION: /note="Identification method S"
; ;
US-07-932-379A-11

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-07-932-379A-11 ..
Align seg 1/1 to: US-07-932-379A-11 from: 1 to: 22

49 GlnThrGlyPheValArgTyr 55
|||||
2 CAAACCGCGTTTGACAGATAT 22

seq_name: /cgn2_6/ptodata/2/1na/5A.COMB.seq:US-07-932-379A-12

seq_documentation_block:
; Sequence 12, Application US/07932379A
; Patent No. 5468852
; GENERAL INFORMATION:
; APPLICANT: Ohashi, Tetsuo
; APPLICANT: Toda, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ozaki, Hiroko
; APPLICANT: Nishimura, Nasyuki
; APPLICANT: Shirasaki, Yoshinari
; APPLICANT: Yamagata, Koichi
; TITLE OF INVENTION: Oligonucleotides for Detecting
; TITLE OF INVENTION: Bacteria and Detection Method
; TITLE OF INVENTION: Using Same
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/932,379A
; FILING DATE: 19920819
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 1327-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli H10407
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..22
; OTHER INFORMATION: /label=oligonucleotide
; ;
; OTHER INFORMATION: /note="Identification method S"
; ;
US-07-932-379A-12

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-07-932-379A-12/rev ..

Align seg 1/1 to reverse of: US-07-932-379A-12 from: 1 to: 22

228 SerGIuValAspIleTYrAsn 234

21 TCAGAGCTTGACATATATATAC 1

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-379-295-11

seq_documentation_block:

; Sequence 11, Application US/08379295

; Patent No. 5516898

; GENERAL INFORMATION:

; APPLICANT: Ohashi, Tetsuo

; APPLICANT: Toda, Jun

; APPLICANT: Fukushima, Shigeru

; APPLICANT: Ozaki, Hiroko

; APPLICANT: Nishimura, Nasyuki

; APPLICANT: Yamagata, Koichi

; TITLE OF INVENTION: Oligonucleotides for Detecting

; TITLE OF INVENTION: Bacteria and Detection Method

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Birch, Stewart, Kolasch & Birch

; STREET: 301 N. Washington St.

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22046-3487

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/379,295

; FILING DATE:

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/932,379A

; FILING DATE: 19-AUG-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Raymond C.

; REGISTRATION NUMBER: 21,066

; REFERENCE/DOCKET NUMBER: 1327-106P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-241-1300

; TELEFAX: 703-241-2848

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli H10407

; FEATURE:

; NAME/KEY: -

; LOCATION: 1..22

; OTHER INFORMATION: /label= oligonucleotide

; OTHER INFORMATION: /note= "identification method S"

; US-08-379-295-11

alignment_scores:

Quality: 7.00

Ratio: 1.000

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-379-295-11 ..

Align seg 1/1 to: US-08-379-295-11 from: 1 to: 22

49 GlnThrGlyPheValArgTyr 55

2 CAACCGCGCTTGTCAGATAT 22

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-379-295-12

seq_documentation_block:

; Sequence 12, Application US/08379295

; Patent No. 5516898

; GENERAL INFORMATION:

; APPLICANT: Ohashi, Tetsuo

; APPLICANT: Toda, Jun

; APPLICANT: Fukushima, Shigeru

; APPLICANT: Ozaki, Hiroko

; APPLICANT: Nishimura, Nasyuki

; APPLICANT: Yamagata, Koichi

; TITLE OF INVENTION: Oligonucleotides for Detecting

; TITLE OF INVENTION: Bacteria and Detection Method

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Birch, Stewart, Kolasch & Birch

; STREET: 301 N. Washington St.

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22046-3487

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/379,295

; FILING DATE:

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/932,379A

; FILING DATE: 19-AUG-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Raymond C.

; REGISTRATION NUMBER: 21,066

; REFERENCE/DOCKET NUMBER: 1327-106P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-241-1300

; TELEFAX: 703-241-2848

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Escherichia coli H10407

; FEATURE:

; NAME/KEY: -

; LOCATION: 1..22

; OTHER INFORMATION: /label= oligonucleotide

; OTHER INFORMATION: /note= "identification method S"

; US-08-379-295-11


```
OTHER INFORMATION: /note="Identification method S"
US-08-379-295-12

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-379-295-12/rev ..

Align seq 1/1 to reverse of: US-08-379-295-12 from: 1 to: 22

228 SerGIuValasPIeTYFRsn 234
|||||
21 TCAGAGGTTGACATATATAC 1

seq_name:*/cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-379-296-11

seq_documentation_block:
; Sequence 11, Application US/08379296
; Patent No. 5525718
; GENERAL INFORMATION:
; APPLICANT: Ohashi, Tetsuo
; APPLICANT: Toda, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ozaki, Hiroko
; APPLICANT: Nishimura, Nasyuki
; APPLICANT: Shirasaki, Yoshinari
; APPLICANT: Yamagata, Koichi
; TITLE OF INVENTION: Oligonucleotides for Detecting Bacteria
; TITLE OF INVENTION: and Detection Method Using Same
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,296
; FILING DATE: 27-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,379
; FILING DATE: 19-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 2036-102P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli H10407
; FEATURE:
; NAME/KEY:
```

```
LOCATION: 1..22
OTHER INFORMATION: /label=oligonucleotide
OTHER INFORMATION: /note="Identification method S"
US-08-379-296-11

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-379-296-11 ..

Align seq 1/1 to: US-08-379-296-11 from: 1 to: 22

49 GlnThrGlyPheValArgTYr 55
|||||
2 CAAACCGGCTTGTGCAGATAT 22

seq_name:*/cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-379-296-12

seq_documentation_block:
; Sequence 12, Application US/08379296
; Patent No. 5525718
; GENERAL INFORMATION:
; APPLICANT: Ohashi, Tetsuo
; APPLICANT: Toda, Jun
; APPLICANT: Fukushima, Shigeru
; APPLICANT: Ozaki, Hiroko
; APPLICANT: Nishimura, Nasyuki
; APPLICANT: Shirasaki, Yoshinari
; APPLICANT: Yamagata, Koichi
; TITLE OF INVENTION: Oligonucleotides for Detecting Bacteria
; TITLE OF INVENTION: and Detection Method Using Same
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,296
; FILING DATE: 27-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,379
; FILING DATE: 19-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 2036-102P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli H10407
```


FEATURE:
NAME/KEY: -
LOCATION: 1.22
OTHER INFORMATION: /label=oligonucleotide
OTHER INFORMATION: /note="identification method S"
US-08-379-296-12

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:

US-09-528-682-1 x US-08-379-296-12/rev ..

Align seg 1/1 to reverse of: US-08-379-296-12 from: 1 to: 22

228 sergluvalaspietyrasn 234

|||||
21 TCACAGCTGCATATATAC 1

seq_name: /cgn2.6/ptodata/2/1na/5A.COMB.seq:US-08-328-710A-26

seq_documentation_block:

Sequence 26, Application US/08328710A

Patent No. 5795717

GENERAL INFORMATION:

APPLICANT: Nakayama, Tomoko

APPLICANT: Tada, Jun

APPLICANT: Fukushima, Shigeru

APPLICANT: Ohashi, Tetsuo

TITLE OF INVENTION: Oligonucleotides for detecting bacteria and detection

TITLE OF INVENTION: process

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch and Birch

STREET: PO Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/328,710A

FILING DATE: 25-OCT-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Weiner, Marc S

REGISTRATION NUMBER: 32,181

REFERENCE/DOCKET NUMBER: 1422-202P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX: 248345

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 22 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Vibrio cholerae

US-08-328-710A-26

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:

US-09-528-682-1 x US-08-328-710A-26/rev ..

Align seg 1/1 to reverse of: US-08-328-710A-26 from: 1 to: 22

122 sergluvalaspietyrasn 128

|||||
21 TCCCAATATATGATGCTAT 1

seq_name: /cgn2.6/ptodata/2/1na/6B.COMB.seq:US-09-311-260-134

seq_documentation_block:

Sequence 134, Application US/09311260

Patent No. 6214555

GENERAL INFORMATION:

APPLICANT: Leushner, James

APPLICANT: Hul, May

APPLICANT: Dunn, James M.

TITLE OF INVENTION: METHOD, COMPOSITIONS AND KIT FOR DETECTION OF

TITLE OF INVENTION: MICROORGANISMS AND BI-DIRECTIONAL SEQUENCING OF NUCLEIC ACI

TITLE OF INVENTION: POLYMERS

NUMBER OF SEQUENCES: 189

CORRESPONDENCE ADDRESS:

ADDRESSEE: Oppehl & Larson LLP

STREET: P.O. Box 5270

CITY: Frisco

STATE: CO

COUNTRY: US

ZIP: 80443-5270

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb storage

OPERATING SYSTEM: IBM compatible

SOFTWARE: Word Perfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/311,260

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Larson, Marina T.

REGISTRATION NUMBER: 32,038

REFERENCE/DOCKET NUMBER: YGEN.P-058-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (970) 668-2050

TELEFAX: (970) 668-2082

TELEX:

INFORMATION FOR SEQ ID NO: 134:

SEQUENCE CHARACTERISTICS:

LENGTH: 22

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

ANTI-SENSE: yes

FRAGMENT TYPE: Internal

US-09-311-260-134

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-311-260-134

Align seg 1/1 to: US-09-311-260-134 from: 1 to: 22

7 ArglaasSerArgProPro 13
|||||
1 CGGGCAGATCTGACCTCTCT 21

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-968-046-26

seq_documentation_block:

; Sequence 26, Application US/08968046
; Patent No. 6218110

; GENERAL INFORMATION:

; APPLICANT: Nakayama, Tomoko

; APPLICANT: Tada, Jun

; APPLICANT: Fukushima, Shigeru

; APPLICANT: Ohashi, Tetsuo

; TITLE OF INVENTION: Oligonucleotides for detecting bacteria

; TITLE OF INVENTION: and detection process

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Birch, Stewart, Kolasch and Birch

; STREET: PO Box 747

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22040-0747

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/968, 046

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/328, 710

; FILING DATE: 25-OCT-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Weiner, Marc S

; REGISTRATION NUMBER: 32,181

; REFERENCE/DOCKET NUMBER: 1422-202P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 205-8000

; TELEFAX: (703) 205-8050

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 26:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Vibrio cholerae

; US-08-968-046-26

alignment_scores:

Quality: 7.00 Length: 7

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-968-046-26/rev

Align seg 1/1 to reverse of: US-08-968-046-26 from: 1 to: 22

122 SerGlnIleTyrGlyTyr 128
|||||
21 TCCCAATATATGATGCTAT 1

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-363-585-12

seq_documentation_block:

; Sequence 12, Application US/08363585

; Patent No. 5683872

; GENERAL INFORMATION:

; APPLICANT: Rudert, William A.

; APPLICANT: Trucco, Massimo

; TITLE OF INVENTION: Polymers of Oligonucleotide probes

; TITLE OF INVENTION: As The Bound Ligands For Use In Reverse

; NUMBER OF SEQUENCES: 112

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: University of Pittsburgh

; STREET: Office of Intellectual Property

; CITY: Pittsburgh

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 15260

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5-1/4" low density diskette

; COMPUTER: IBM PC or compatibles

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/363,585

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/786,228

; FILING DATE: 31-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Frederick H. Coleman; Mary-Elizabeth Buckles

; REGISTRATION NUMBER: 28,061; 31,907

; REFERENCE/DOCKET NUMBER: 92-232

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 412/288-4164

; TELEFAX: 412/288-3063

; TELEX: 277871

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 27 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: synthetic DNA

; PUBLICATION INFORMATION:

; AUTHORS: Rudert, W.A.

; AUTHORS: Trucco, M.

; TITLE: A No. 5683872el Approach to Rapid HLA Class II

; JOURNAL: Molecular Typing

; VOLUME: 2

; PAGES: 352-356

; DATE: 1992

; RELEVANT RESIDUES IN SEQ ID NO: 12: 1 to 27

; US-08-363-585-12

alignment_scores:

Quality: 7.00 Length: 7

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-363-585-12

Align seg 1/1 to: US-08-363-585-12 from: 1 to: 27

18 ArgSerGlyGlyLeuMetPro 24
|||||
5 AGATCTGGCGGCTGATCCG 25

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-09-142-355B-11

seq_documentation_block:
; Sequence 11, Application US/09142355B
; Patent No. 6184012
; GENERAL INFORMATION:
; APPLICANT: Neri, Dario
; APPLICANT: Demattis, Salvatore
; APPLICANT: Huber, Adrain
; APPLICANT: Viti, Francesca
; APPLICANT: Tawfik, Dan. S.
; APPLICANT: Winter, Gregory Paul
; TITLE OF INVENTION: Isolation of Enzymes
; FILE REFERENCE: 2224/0665
; CURRENT APPLICATION NUMBER: US/09/142,355B
; CURRENT FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: GB 9608540.2
; PRIOR FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/GB97/01153
; PRIOR FILING DATE: 1997-04-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: primer
US-09-142-355B-11

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-142-355B-11/rev ..

Align seg 1/1 to reverse of: US-09-142-355B-11 from: 1 to: 42

61 SerThrSerLeuSerLeuArg 67
|||||
35 AGTACACGCTCAGCTCCG 15

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-276-452A-65

seq_documentation_block:
; Sequence 65, Application US/08276452A
; Patent No. 5646029
; GENERAL INFORMATION:
; APPLICANT: Chen, Chao-Guang
; APPLICANT: Mau, Shiao-Lim
; APPLICANT: Du, He
; APPLICANT: Gane, Alison M
; APPLICANT: Bacic, Antony
; APPLICANT: Clarke, Adrienne E
; TITLE OF INVENTION: Plant Arabidnolactan Protein (AGP) Genes
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: United States of America
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,452A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 27-91A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; TELEX: 49617824

INFORMATION FOR SEQ ID NO: 65:

SEQUENCE CHARACTERISTICS:

LENGTH: 85 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: misc_feature

LOCATION: 1..19

OTHER INFORMATION: /note= "T7 promoter sequence"

NAME/KEY: misc_feature

LOCATION: 68..85

OTHER INFORMATION: /note= "Adaptor sequence is equal

to nucleotides 444-461 of the NaAgp1 cDNA"

US-08-276-452A-65

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-276-452A-65 ..

Align seg 1/1 to: US-08-276-452A-65 from: 1 to: 85

211 LeuArgGluTyGlnSerLys 217
|||||
64 TTGCGGGAGTATCAGTCATAA 84

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-798-744-65

seq_documentation_block:
; Sequence 65, Application US/08798744
; Patent No. 5830747
; GENERAL INFORMATION:
; APPLICANT: Chen, Chao-Guang
; APPLICANT: Mau, Shiao-Lim
; APPLICANT: Du, He
; APPLICANT: Gane, Alison M
; APPLICANT: Bacic, Antony
; APPLICANT: Clarke, Adrienne E
; TITLE OF INVENTION: Plant Arabidnolactan Protein (AGP) Genes
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: United States of America
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS


```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/798,744
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/276,452
; FILING DATE: 18-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 27-91A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303)499-8080
; TELEFAX: (303)499-8089
; TELEX: 49617824
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..19
; OTHER INFORMATION: /note= "T7 promoter sequence"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 68..85
; OTHER INFORMATION: /note= "Adaptor sequence is equal
; OTHER INFORMATION: to nucleotides 444-461 of the NaAgp1 cDNA"
; US-08-798-744-65

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-798-744-65 ..

Align seg 1/1 to: US-08-798-744-65 from: 1 to: 85

211 LeuArgGluTyrGlnSerIys 217
|||||
64 TTGCGGAGATCACTCAAAA 84

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-060-756-266
documentation_block:
; sequence 266, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/09/060,756
; CURRENT FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 266
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: unsure
```

```

; LOCATION: (various positions within the sequence)
; OTHER INFORMATION: applicants are uncertain of bases designated as "n"
; US-09-060-756-266

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-060-756-266 ..

Align seg 1/1 to: US-09-060-756-266 from: 1 to: 217

63 SerLeuSerLeuArgSerAla 69
|||||
10 TCCATTATCGCTCCGCTCTGCA 30

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-780C-95
documentation_block:
; sequence 95, Application US/08471780C
; Patent No. 5759808
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,780C
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/106,944
; FILING DATE: 17-AUG-1993
; APPLICATION NUMBER: FR 92402326.0
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 93401310.3
; FILING DATE: 21-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Potler, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 04958.0008-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-471-780C-95

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
```


Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-471-780C-95/rev ..

Align seg 1/1 to reverse of: US-08-471-780C-95 from: 1 to: 433

62 ThrsleuserleuAgsr 68
|||||
340 ACCTCCTGTGCTAGGTG 320

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-467-282B-95

seq_documentation_block:

; Sequence 95, Application US/08467282B
; Patent No. 5800988

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESS:

ADDRESS: Flinagan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467, 282B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/106, 944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958, 0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 95:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-467-282B-95

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x US-08-467-282B-95/rev ..

Align seg 1/1 to reverse of: US-08-467-282B-95 from: 1 to: 433

62 ThrsleuserleuAgsr 68

|||||
340 ACCTCCTGTGCTAGGTG 320

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-471-282A-95

seq_documentation_block:

; Sequence 95, Application US/08471282A

; Patent No. 5840853

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESS:

ADDRESS: Flinagan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471, 282A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/106, 944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958, 0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 95:

SEQUENCE CHARACTERISTICS:

LENGTH: 433 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-471-282A-95

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x US-08-471-282A-95/rev ..

Align seg 1/1 to reverse of: US-08-471-282A-95 from: 1 to: 433

62 ThrsleuserleuAgsr 68
|||||

340 ACCTCCTGTGCTAGGTG 320

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-466-710C-95

seq_documentation_block:

; Sequence 95, Application US/08466710C

; Patent No. 5874541


```
/
/ GENERAL INFORMATION:
/ APPLICANT: Casterman, Cecile
/ APPLICANT: Hamers, Raymond
/ TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
/ NUMBER OF SEQUENCES: 130
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
/ STREET: 1300 I Street, N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20005-3315
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/466,710C
/ FILING DATE:
/ CLASSIFICATION:
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/106,944
/ FILING DATE: 17-AUG-1993
/ APPLICATION NUMBER: FR 92402326.0
/ FILING DATE: 21-AUG-1992
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 93401310.3
/ FILING DATE: 21-MAY-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Potter, Jane E.R.
/ REGISTRATION NUMBER: 33,332
/ REFERENCE/DOCKET NUMBER: 04958.0008-00000
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/
/ INFORMATION FOR SEQ ID NO: 95:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 433 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: CDNA
/
/ US-08-466-710C-95
/
/ alignment_scores:
/      Quality: 7.00      Length: 7
/      Ratio: 1.000      Gaps: 0
/      Percent Similarity: 100.000      Percent Identity: 100.000
/
/ ment_block:
/ 09-528-682-1 x US-08-466-710C-95/rev ..
/
/ Align seg 1/1 to reverse of: US-08-466-710C-95 from: 1 to: 433
/
/ 62 ThrsSerLeuSerLeuArgSer 68
/ |||||||||||||||||||
/ 340 ACCTCCTGTGCGCTAAGTGG 320
/
/ seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-471-284B-95
/
/ seq_documentation_block:
/ ; Sequence 95, Application US/08471284B
/ ; Patent No. 6005079
/
/ GENERAL INFORMATION:
/ APPLICANT: Casterman, Cecile
/ APPLICANT: Hamers, Raymond
/ TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
/ NUMBER OF SEQUENCES: 130
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
/ STREET: 1300 I Street, N.W.
```

```
/
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20005-3315
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/471,284B
/ FILING DATE:
/ CLASSIFICATION: 530
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/106,944
/ FILING DATE: 17-AUG-1993
/ APPLICATION NUMBER: FR 92402326.0
/ FILING DATE: 21-AUG-1992
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 93401310.3
/ FILING DATE: 21-MAY-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Potter, Jane E.R.
/ REGISTRATION NUMBER: 33,332
/ REFERENCE/DOCKET NUMBER: 04958.0008-00000
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/
/ INFORMATION FOR SEQ ID NO: 95:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 433 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: CDNA
/
/ US-08-471-284B-95
/
/ alignment_scores:
/      Quality: 7.00      Length: 7
/      Ratio: 1.000      Gaps: 0
/      Percent Similarity: 100.000      Percent Identity: 100.000
/
/ alignment_block:
/ US-09-528-682-1 x US-08-471-284B-95/rev ..
/
/ Align seg 1/1 to reverse of: US-08-471-284B-95 from: 1 to: 433
/
/ 62 ThrsSerLeuSerLeuArgSer 68
/ |||||||||||||||||||
/ 340 ACCTCCTGTGCGCTAAGTGG 320
/
/ seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-468-739C-95
/
/ seq_documentation_block:
/ ; Sequence 95, Application US/08468739C
/ ; Patent No. 6015695
/
/ GENERAL INFORMATION:
/ APPLICANT: Casterman, Cecile
/ APPLICANT: Hamers, Raymond
/ TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
/ NUMBER OF SEQUENCES: 130
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
/ STREET: 1300 I Street, N.W.
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20005-3315
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
```


SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,739C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-468-739C-95

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-468-739C-95/rev ..

Align seg 1/1 to reverse of: US-08-468-739C-95 from: 1 to: 433

62 ThrsEulerSerLeuArgSer 68
|||||
340 ACGTCCTGTGCTAAGTCG 320

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-471-780C-96
seq_documentation_block:
Sequence 96, Application US/08471780C
Patent No. 5759808
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,780C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-471-780C-96

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-471-780C-96/rev ..

Align seg 1/1 to reverse of: US-08-471-780C-96 from: 1 to: 449

62 ThrsEulerSerLeuArgSer 68
|||||
352 ACGTCCTGTGCTAAGTCG 332

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-467-282B-96

seq_documentation_block:
Sequence 96, Application US/08467282B
Patent No. 5800988
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,282B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-467-282B-96

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-467-282B-96/rev ..

Align seg 1/1 to reverse of: US-08-467-282B-96 from: 1 to: 449

62 ThrsEulerleuArgSer 68
|||||
352 ACGTCCTGTCGCTAAGCTCG 332

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-282A-96

seq_documentation_block:
Sequence 96, Application US/08471282A
Patent No. 5840853
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,282A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-471-282A-96

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-471-282A-96/rev ..

Align seg 1/1 to reverse of: US-08-471-282A-96 from: 1 to: 449

62 ThrsEulerleuArgSer 68
|||||
352 ACGTCCTGTCGCTAAGCTCG 332

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-466-710C-96

seq_documentation_block:
Sequence 96, Application US/08466710C
Patent No. 5874541
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,710C
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-466-710C-96

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-466-710C-96/rev ..

Align seg 1/1 to reverse of: US-08-466-710C-96 from: 1 to: 449

62 ThrsEulerSerLeuArgSer 68
|||||
352 ACCTCTTGTCTAGCTCG 332

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-471-284B-96

seq_documentation_block:

; Sequence 96, Application US/08471284B
; Patent No. 6005079

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Flinagan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,284B

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/106,944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potier, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958, 0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 96:

SEQUENCE CHARACTERISTICS:

LENGTH: 449 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-471-284B-96

alignment_scores:

Quality: 7.00 Length: 7

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-471-284B-96/rev ..

Align seg 1/1 to reverse of: US-08-471-284B-96 from: 1 to: 449

62 ThrsEulerSerLeuArgSer 68
|||||
352 ACCTCTTGTCTAGCTCG 332

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-468-739C-96

seq_documentation_block:

; Sequence 96, Application US/08468739C

; Patent No. 6015695

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Flinagan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,739C

FILING DATE: 06-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/106,944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potier, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958, 0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 96:

SEQUENCE CHARACTERISTICS:

LENGTH: 449 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-468-739C-96

alignment_scores:

Quality: 7.00 Length: 7

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-08-468-739C-96/rev ..

Align seg 1/1 to reverse of: US-08-468-739C-96 from: 1 to: 449

62 ThrsEulerSerLeuArgSer 68
|||||
352 ACCTCTTGTCTAGCTCG 332

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-050-259B-17

seq_documentation_block:

; Sequence 17, Application US/08050259B


```
; Patent No. 6127169
; GENERAL INFORMATION:
; APPLICANT: Livio Mallucci
; APPLICANT: Valerie Wells
; TITLE OF INVENTION: CELL GROWTH INHIBITORS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; City: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,259B
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB91/01898
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9023907.0
; FILING DATE: 02-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreyer, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-58324/WHD/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-050-259B-17

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

nment_block:
09-528-682-1 x US-08-050-259B-17/rev ..

Align seg 1/1 to reverse of: US-08-050-259B-17 from: 1 to: 497

161 GYTYRARGLEALAGLYPHE 167
|||||
446 GGCTACAGCGTGGCTGCTC 426

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-770-544-19

seq_documentation_block:
; Sequence 19, Application US/08770544
; Patent No. 5907085
; GENERAL INFORMATION:
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Ling, Kai-Shu
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
; TITLE OF INVENTION: THEIR USES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
```

```
; STREET: Clinton Square, P.O. Box 1051
; City: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,544
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60009008
; FILING DATE: 21-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-770-544-19

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-770-544-19/rev ..

Align seg 1/1 to reverse of: US-08-770-544-19 from: 1 to: 534

59 TYRVALSERTHRSEULEUSER 65
|||||
501 TACITTCACACTCTGTCT 481

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-133-804-3

seq_documentation_block:
; Sequence 3, Application US/08133804
; Patent No. 5534254
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```



```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/133,804
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Kelley, Robin D.
: REGISTRATION NUMBER: 34,637
: REFERENCE/DOCKET NUMBER: 2054/22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-248-7477
: TELEFAX: 617-248-7100
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 779 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..758
: OTHER INFORMATION: /product= "26-10 sfv" with
: OTHER INFORMATION: C-terminal Gly4-Cys"
:
: US-08-133-804-3

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-133-804-3  ..

Align seg 1/1 to: US-08-133-804-3 from: 1 to: 779

15 GlutIeLySArGSeRgLyGly 21
|||||
729 GAGATTAAACGTCGCGGGCA 749

seq_name: /cgn2.6/prodata/2/lna/5A.COMB.seq:US-08-461-838-3

seq_documentation_block:
: Sequence 3, Application US/08461838
: Patent No. 5753204
: GENERAL INFORMATION:
: APPLICANT: Huston, James S.
: APPLICANT: Oppermann, Hermann
: APPLICANT: Houston, L. L.
: APPLICANT: Rling, David B.
: TITLE OF INVENTION: Blosynthetic Binding Proteins For
: TITLE OF INVENTION: Imaging
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
: STREET: Exchange Place, 53 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461,838
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Kelley, Robin D.
: REGISTRATION NUMBER: 34,637
: REFERENCE/DOCKET NUMBER: 2054/22
```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-248-7477
: TELEFAX: 617-248-7100
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 779 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..758
: OTHER INFORMATION: /product= "26-10 sfv" with
: OTHER INFORMATION: C-terminal Gly4-Cys"
:
: US-08-461-838-3

alignment_scores:
      Quality: 7.00      Length: 7
      Ratio: 1.000      Gaps: 0
      Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-461-838-3  ..

Align seg 1/1 to: US-08-461-838-3 from: 1 to: 779

15 GlutIeLySArGSeRgLyGly 21
|||||
729 GAGATTAAACGTCGCGGGCA 749
```


us-09-528-682-1.01ip2n.rml

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,045C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/694,733
FILING DATE: 02-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/271,222
FILING DATE: 06-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196C
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 777 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..777
US-08-449-045C-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-449-045C-1 ..
US-09-528-682-1 x US-08-449-045C-1

Align seg 1/1 to: US-08-449-045C-1 from: 1 to: 777

139 GAGTACTTTGACGACGAGTACTCAATGATATCAACCTTTATGATCATGTC 188
45 AARGLYTHRGINLTHRGILYPHEVALRG 54
189 AAGAGGACTCAGACGGGATTTGTTAGG 216

seq_name: /cgn2_6/ptodata/2/line/backfiles1.seq:5223610-1
seq_documentation_block:
Patent No. 5223610
APPLICANT: Burton, Frank H.; Sutcliffe, Gregor
TITLE OF INVENTION: CHOLERA TOXIN GENE REGULATED BY GROWTH
HORMONE PROMOTER
NUMBER OF SEQUENCES: 18
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/528,852
FILING DATE: 18-MAY-1990
SEQ ID NO: 1
LENGTH: 2020
5223610-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x 5223610-1 ..

```


Date: Jun 18, 2002 7:26 PM

About: Results were produced
Copyright (c) 1993-2

Command line parameters:

```
-O -cgn1.1/USPTO.spool/US09578683/runa_c18062002.082443.7808/app.query.fasta_1.6688
-DB=genEmul1 -QFMT=fastest -SUFFIX=p2n.ire -GAP=12.000
-GAEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -YGAPOP=10.000 -YGAEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -XGAPOP=10.000 -XGAEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=1000 -DOCALLIGN=200 -THR_SCORE=pct
-TR_MAX=100 -THR_MIN=0 -ALIGN=50 -MOD=LOCAL -OUTFMT=ps
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09528682.@cgn1.1.8225 -NCPU=6 -ICPU=3 -LONGLOG
-DEXTIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPUX -WAIT -THREADS=1
```

Search information block:

Query length: 240

Database sequences: 1797656

Search time (sec): 3690.420000

score_list:

Sequence	Strd	Orig	Len	Documentation
gb_Da:AF242417	+	1304..00	2281..18	1..e-118 777
gb_Da:AF242417	+	1304..00	2281..18	1..e-118 777
gb_Da:P207LTA	+	1304..00	2281..18	1..e-118 777
gb_Da:ENSLTPA	+	1304..00	2280..15	1..e-118 777
gb_Da:ENSLTPA	+	1304..00	2280..15	1..e-118 866
gb_Pat:R04315	+	1300..00	2279..80	1..3e-118 777
gb_Pat:R04315	+	1300..00	2279..80	1..3e-118 898
gb_Pat:R00613	+	1298..00	2270..63	1..2e-118 777
gb_Pat:EMD013	+	1293..00	2262..52	1..e-118 773
gb_Pat:R03421	+	1292..00	2256..36	2..6e-117 1148
gb_Da:AB011677	+	1290..00	2250..53	5..3e-117 1434
gb_Da:ECOT0XA	+	1287..00	2249..73	6..2e-117 934
gb_Pat:R00731	+	1287..00	2246..57	9..0e-117 1275
gb_Pat:R16149	+	1183..00	2065..25	6..8e-107 711
gb_Pat:AR118597	+	1183..00	2065..25	6..8e-107 711
gb_Da:EC0LTA	+	1183..00	2065..35	8..7e-107 868
gb_Da:AR455584	+	1093..00	1910..14	4..9e-98 777
gb_Da:AR463400	+	1093..00	1910..14	4..9e-98 777
gb_Da:AR463401	+	1093..00	1910..14	4..9e-98 777
gb_Da:VIBCT0XABA	+	1093..00	1904..72	1..9e-98 1372
gb_Pat:R16422	+	1088..00	1902..03	1..4e-97 723
gb_Pat:AR118598	+	1088..00	1902..03	1..4e-97 723
gb_Pat:R00132	+	1088..00	1897..63	2..5e-97 1148
gb_Da:VCT0XAB21	+	1088..00	1896..00	3..0e-97 1362
gb_Da:VIBCT0XAB5	+	1088..00	1895..95	3..0e-97 1369
gb_Da:VCT0XAB56	+	1088..00	1895..76	3..1e-97 1397
gb_Pat:VCC0TX	+	1088..00	1892..25	4..9e-97 2020
gb_Pat:R03512	+	1088..00	1892..25	4..9e-97 2020
gb_Da:AR1175708	+	1088..00	1887..07	9..5e-97 3480
gb_Pat:AR414364	+	1088..00	1885..41	1..2e-96 4275
gb_Da:AB004224	+	984..00	1630..53	3..7e-96 10646
gb_Da:VIBCT0XA	+	984..00	1630..53	1..9e-82 777
gb_Da:ECOE0XOHL	+	632..00	1182..78	1..6e-57 1262
gb_Da:ECOE0LTA	+	679..50	1178..86	2..7e-57 1200
gb_Da:ECOE0LT	+	357..00	615..92	6..1e-26 774
gb_Da:EC0L757	+	340..00	568..53	2..0e-24 595
gb_Da:ECOE0LTAB2	+	247..00	437..20	5..5e-16 861
gb_Da:VCC0TXAB2	+	182..00	307..17	9..5e-09 165
gb_Da:VCC0TXAB	+	182..00	306..35	1..1e-08 938
gb_Pat:R02702	+	166..00	279..79	3..2e-07 795
gb_Pat:R02702	+	166..00	279..79	3..2e-07 795
AF242417	+	Escherichia coli	heat-labile enterotoxin	777
P207LTA	+	Plasmid P307 (from E.coli)	heat-labile enterotoxin	777
ENSLTPA	+	E.coli LTA gene for labile enterotoxin		777
ENSLTPA	+	Plasmid ENT P307 (enteric enterotoxin)		866
R04315	+	Plasmid ENT-R pC586 (enteric enterotoxin)		777
R04315	+	Artificial sequence LTA		898
R00613	+	DNA encoding a component of enteric enterotoxin		773
EMD013	+	Plasmid EMD 299 (from E.coli)		773
R03421	+	DNA encoding Lth of enteric enterotoxin		1148
AB011677	+	Escherichia coli gene encoding enteric enterotoxin		1434
ECOT0XA	+	E.coli toxin gene encoding enteric enterotoxin		934
R00731	+	heat-labile enterotoxin		1275
R16149	+	A subunit of a labile toxin		711
AR118597	+	A subunit of a labile toxin		711
EC0LTA	+	Sequence 5 from patent		868
AR455584	+	col gene elta encoded by		777
AR455584	+	Vibrio cholerae strain		777
AR463400	+	Vibrio cholerae strain		777
AR463401	+	Vibrio cholerae strain		777
VIBCT0XABA	+	Vibrio cholerae strain		1372
R16422	+	Vibrio cholerae toxin		723
AR118598	+	A subunit of cholera toxin		723
R00132	+	Sequence 7 from patent		1148
VCT0XAB21	+	DNA coding of cholera toxin		1362
VIBCT0XAB5	+	X58736 DNA coding of cholera toxin		1369
VCT0XAB56	+	D30053 Vibrio cholerae toxin		1397
VCC0TX	+	D35785 V. cholerae genes for toxin		2020
R03512	+	X00171 Vibrio cholera toxin		

gb_pat:AA06082	163.00	274.56	6.3e-07	791	! A060820 PERK620 HindII-fragment
gb_pat:AB06083	157.00	274.56	6.3e-07	791	! A060830 PERK620 HindII-fragment
gb_ba:VIBT08	157.00	268.15	1.4e-06	512	! K01170 Vibrio cholerae toxA
gb_ba:VIBCTA2	117.00	192.45	2.9e-06	251	! J01847 Vibrio cholerae cholerae
gb_ba:ECOPHIEG5	117.00	192.45	0.0234	901	! M61015 Escherichia coli heatlab
gb_pat:BA69697	109.50	166.72	0.4919	2724	! BA69697 Sequence 15 from Pat
gb_pat:BD009884	109.50	166.72	0.4919	2724	! BD009884 Recombinant toxin
gb_ba:VIBCTA1	108.00	168.62	0.0383	255	! J01846 Vibrio cholerae cholerae
gb_pat:AC103425	106.00	197.14	0.0128	72	! J03425 Sequence 1 from Patentre
gb_hcg:AC024743	101.00	114.55	511.65	167751	! AC107423 Homo sapiens cirm
gb_hcg:AC028396	101.00	113.27	603.19	191963	! AC024537 Homo sapiens cirm
gb_in:AC078796	99.50	157.85	1.98	1344	! AF078796 Caenorhabditis elate
gb_pat:AX020906	99.50	157.85	2.19	1344	! AX020906 Sequence 2 from Pat
gb_in:AF114440	97.00	120.33	243.66	1461	! AF114440 Caenorhabditis elate
gb_ba:BEPTCPSB	96.00	140.20	19.07	4501	! AL049485 Streptomyces coelic
gb_ba:AE001886	96.00	131.82	55.86	10854	! M14493 Bordetella pertussis
gb_ba:AE0089105	96.00	120.99	70.44	13124	! AE001886 Deinococcus radiat
gb_ba:AE008070	96.00	114.41	520.84	13147	! AE0089105 Agrobacterium tum
gb_pl:AE003314	96.00	114.41	520.84	67580	! AE008070 Agrobacterium tum
gb_ba:MYV016	94.50	111.97	551.27	53662	! AF003314 Oryza sativa geno
gb_in:CONSOT016	94.50	111.97	2.7e+03	194439	! AL021841 Mycobacterium tubc
gb_in:AB030305	94.00	137.87	25.72	3973	! AL590443 chromosome III c
gb_pat:ES1095	94.00	137.87	25.72	3973	! AB030305 Pteris rapae mRNA
gb_ba:BBE245366	93.00	151.24	4.62	810	! ES1095 Apoptosis-inducing p
gb_ba:BBE245367	93.00	151.24	4.62	810	! AJ245366 Bordetella pertussis
gb_ba:BBE245367	93.00	147.57	7.41	1192	! AJ245367 Bordetella pertussis
gb_ba:BBA006155	93.00	136.74	29.70	3716	! AJ006153 Bordetella pertussis
gb_ba:BBA006153	93.00	147.47	7.50	1204	! AJ006153 Bordetella pertussis
gb_ba:BBA006157	93.00	147.02	7.95	1262	! AJ006157 Bordetella pertussis
gb_ba:BBE7363	93.00	147.02	7.95	1263	! A007363 Bordetella pertussis
gb_ba:BBE7364	93.00	147.02	7.95	1263	! A007364 Bordetella pertussis
gb_pat:103380	93.00	143.03	13.25	1919	! I03380 Sequence 5 from Patate
gb_in:AB037676	93.00	136.74	29.70	3716	! AB037676 Pteris brassicae m
gb_pat:ES1096	93.00	136.74	29.70	3716	! ES1096 Apoptosis-inducing p
gb_ba:BBPT0X	93.00	135.56	34.58	4210	! M13223 Bordetella pertussis
gb_pat:ED01352	93.00	134.52	39.52	4696	! ED01352 Genomic DNA encoding
gb_ba:BBEPT0XS	93.00	134.04	42.00	4936	! M14378 Bordetella pertussis
gb_com:RABCYCP2B	92.50	141.55	16.03	2045	! M20858 Rabbit cytochrome P
gb_hcg:AC094670	92.50	106.14	1.5e+03	84424	! AC094670 Rattus norvegicus
gb_pr:HSBDJ92CB	92.50	101.95	2.6e+03	131060	! AL132671 Human DNA sequenc
gb_ba:AL646079	92.50	97.33	4.7e+03	2436	! AL646079 Ralstonia solanum
gb_pr:RANPFR	91.50	128.13	24.87	11596	! L19658 Rattus norvegicus FR
gb_ba:AE007993	90.50	121.52	209.37	11596	! AE007993 Agrobacterium tum
gb_ba:AE009027	90.50	120.29	245.18	13197	! AE009027 Agrobacterium tum
gb_pr:HS49E3	90.50	100.49	3.1e+03	105570	! AL160451 Homo sapiens cirm
gb_hcg:AL590046_1	90.50	100.10	3.3e+03	110000	! AL590046 Homo sapiens cirm
gb_hcg:AL590046_2	90.50	100.10	3.3e+03	110000	! AL590046 Homo sapiens cirm
gb_pr:AL590076	90.50	98.33	3.6e+03	119330	! AL590076 Human DNA sequenc
gb_pr:AP001347	90.50	98.16	4.2e+03	134926	! AP001347 Homo sapiens gene
gb_hcg:AL450318	90.50	89.23	6.9e+03	203803	! AL450318 Homo sapiens cirm
gb_pr:AP001660	90.50	89.36	1.3e+04	340000	! AP001660 Homo sapiens gene
gb_ba:AF177946	89.50	130.07	69.91	3925	! AF177946 Streptomyces coelic

gb-pr:BC000650	88.00	135.73	33.82	1641	BC000650 Homo sapiens, Simla	gb-pat:144694	83.50	125.65	123.23	2061	I44694 Sequence 5 from pate
gb-pr:AK021559	88.00	135.12	36.58	1750	AK021559 Homo sapiens cDNA FL	gb-in:AY051462	83.50	122.86	176.32	2764	AY051462 Drosophila melanog
gb-pr:HS0802252	88.00	128.78	82.49	3407	AL137520 Homo sapiens mRNA: c	gb-ba:BC000642	83.50	119.39	274.88	3976	AY090242 Bacillus cereus merr
gb-ba:BP00065A	88.00	125.25	129.71	4336	ML6492 B.bronchiseptica pertus	gb-ba:BC000642	83.50	119.03	288.14	4133	M60446 Bacillus sphaericus
gb-pat:ARI66753	88.00	118.80	226.52	9717	ARI66753 Sequence 1 from paten	gb-pat:AA4408	83.50	119.03	288.14	4133	AA4498 Sequence 1 from pate
gb-ba:AE008681	88.00	118.37	313.30	10165	AE008681 Rickettsia conorii M	gb-pat:ED05304	83.50	119.03	288.14	4133	E05304 DNA encoding insecti
gb-hg:LMF1C8R31.18	88.00	96.10	5.5e+03	105516	AL512786 Homo sapiens chromo	gb-ba:AE009510	83.50	117.36	356.90	4925	AF397025 Brucella abortus
gb-hg:AC011772	88.00	95.71	5.7e+03	110000	Continuation (19 of 22) of I	gb-ba:AE009510	83.50	117.36	356.90	4925	AF397025 Brucella abortus
gb-hg:AC011772	88.00	94.35	6.8e+03	126851	AC0092815 Homo sapiens chromo	gb-hg:AL158039	83.50	83.59	2.7e+04	171132	AL158039 Homo sapiens chr
gb-hg:AC011772	88.00	92.57	8.6e+03	128891	AC011772 Homo sapiens chromo	gb-pr:AC007376	83.50	82.42	3.2e+04	186330	AC007376 Homo sapiens chr
gb-hg:AC046365	88.00	91.28	1.0e+04	175076	AC046365 Homo sapiens clone	gb-pr:AL157938	83.50	82.42	3.2e+04	197019	AL157938 Homo sapiens chr
gb-pr:AL1596326	88.00	91.11	1.0e+04	178309	AL1596326 Homo sapiens clone	gb-ba:HPW113P4	83.50	78.45	3.2e+04	253431	AL104181 Plasmid DNA sequen
gb-pr:AC0118795	88.00	90.61	1.1e+04	187836	AC0118795 Homo sapiens chromo	gb-ba:HPW113P4	83.50	78.45	3.2e+04	253431	AL104181 Plasmid DNA sequen
gb-ba:AF110332	87.50	132.40	7.97	458	AF110332 Mesorhizobium sp. L09	gb-ro:CR00088A	83.00	134.24	40.94	762	M33958 Chinese hamster prion
gb-pr:HS082009	87.50	132.40	51.83	2123	AF032882 Homo sapiens forkhead	gb-ro:CR00088A	83.00	134.24	40.94	762	M33958 Chinese hamster prion
gb-in:LM00085	87.50	90.63	1.1e+04	170908	AL121582 Human DNA sequence	gb-ba:AF110466	83.00	125.14	131.59	1983	AF110466 Aeromonas hydroph
gb-hg:AC008620	87.50	89.26	1.2e+04	177648	AL359716 Leishmania major F	gb-pr:AF003340	83.00	123.34	165.61	2394	AF003340 Homo sapiens R
gb-ro:AC02236	87.50	89.26	1.2e+04	183399	AC008620 Homo sapiens chromo	gb-ro:AF139572	83.00	123.34	165.61	2394	AF139572 Mus musculus carti
gb-pr:AF137379	87.50	89.12	1.3e+04	198392	AC022236 Mus musculus chromo	gb-in:AY070571	83.00	123.16	169.50	2440	AY070571 Drosophila melanog
gb-pr:AF137379	87.50	89.10	1.3e+04	200799	AF137379 Nephroselmis olivac	gb-ba:AE007516	83.00	107.91	1.2e+03	12119	AE007516 Clostridium aceto
gb-hg:AC009443	87.50	88.97	1.4e+04	203447	AC009443 Mus musculus chromo	gb-ba:AE006951	83.00	105.46	1.6e+03	15670	AE006951 Mycobacterium tub
gb-ba:AE042847	87.00	128.59	84.51	2889	AF242847 Bordetella bronchise	gb-in:U41541	83.00	97.23	5.1e+03	39649	U41541 Caenorhabditis eleg
gb-pr:SP0669	87.00	112.72	66.74	15301	AL035226 S.pombe chromosome 1	gb-ba:SCD6	83.00	96.61	5.1e+03	40969	AL153515 Streptomyces coel
gb-pr:AC019013	87.00	93.60	7.5e+03	114036	AL035226 S.pombe chromosome 1	gb-hg:AC0106613	83.00	86.60	1.4e+04	92118	AC106613 Rattus norvegicus
gb-hg:AC092555	86.50	131.47	1.2e+04	163217	AC092555 Rattus norvegicus d	gb-pr:AC005320	83.00	87.20	2.2e+04	106729	AC005320 Human Chromosome
gb-ba:NM00000	86.50	120.98	224.18	5857	X82637 N.meningitidis pdh gene	gb-pr:AC005320	83.00	84.23	2.5e+04	145825	AC005320 Homo sapiens **
gb-ba:AE002482	86.50	114.14	536.35	11968	AE002482 Neisseria meningitidis	gb-hg:AC021833	83.00	83.18	2.9e+04	162813	AC021833 Homo sapiens chr
gb-ba:AE002482	86.50	85.64	2.1e+04	240050	AL627267 Salmonella enterica	gb-pr:AC068919	83.00	83.09	2.9e+04	164350	AC068919 Homo sapiens chr
gb-pat:AK04032	86.50	82.05	3.3e+04	349980	AK04032 Sequence 111 from F	gb-hg:AC0103528	83.00	82.81	3.0e+04	178119	AC103528 Rattus norvegicus
gb-in:AE007768	86.00	125.89	119.51	2435	AF007768 Choriostoma fumifera	gb-hg:AC0103528	83.00	82.81	3.0e+04	178119	AC103528 Rattus norvegicus
gb-ba:AE023593	86.00	125.89	119.51	3190	AB023293 Bacillus thuringiensis	gb-hg:AC023981	83.00	82.77	3.0e+04	169308	AC023981 Homo sapiens chr
gb-ba:AE023593	86.00	114.25	539.04	10790	AE023293 Pseudomonas aeruginosa	gb-ba:SC0590464	83.00	82.77	3.0e+04	178073	AL5990464 Streptococcus coe
gb-ba:AE009010	86.00	114.25	539.04	10790	AE009010 Agrobacterium tumefaciens	gb-pr:HS0508N22	83.00	80.81	3.8e+04	208642	AL1591137 Homo sapiens chr
gb-ba:AE004646	86.00	113.90	556.34	11244	AE004646 Pseudomonas aeruginosa	gb-hg:AC092425	83.00	80.81	3.8e+04	208642	AL1591137 Homo sapiens chr
gb-ba:AE004646	86.00	83.71	2.7e+04	268200	AE004646 Pseudomonas aeruginosa	gb-pr:AF050588	83.00	76.97	1.04.08	312283	AF050588 Homo sapiens chr
gb-com:AY029774	85.50	138.53	23.63	771	AY029774 Ovis aries prion prote	gb-ba:AR06H00NA	82.50	120.11	250.64	3331	L08502 Drosophila melanog
gb-com:RAB00000	85.50	129.15	78.67	2065	X65046 R. abortus DNA sequence 4	gb-in:DR00000E	82.50	119.61	267.31	3331	L08502 Drosophila melanog
gb-hg:AC094941	85.50	128.81	82.13	2139	X65046 R. abortus DNA sequence 4	gb-pat:170275	82.50	119.61	267.31	3331	L08502 Drosophila melanog
gb-pr:AC005371	85.50	92.79	8.3e+03	94186	AC005371 Rattus norvegicus cl	gb-in:AY047339	82.50	119.61	267.31	3331	L08502 Drosophila melanog
gb-hg:AE004569	85.50	88.54	1.4e+04	142600	AE004569 Oryza sativa chromo	gb-in:AY047339	82.50	119.61	267.31	3331	L08502 Drosophila melanog
gb-hg:AE004569	85.50	88.54	1.4e+04	147185	AE004569 Oryza sativa chromo	gb-ba:AE004668	82.50	107.89	1.2e+03	36774	AE004668 Pseudomonas aerug
gb-ba:AE024772	85.50	88.14	2.7e+04	245050	AE024772 Salmonella enterica	gb-hg:AC019572	82.50	96.46	5.2e+03	36774	AE019572 Drosophila melanog
gb-ba:AE000062	85.50	83.43	2.8e+04	251700	AE000062 Aeropyrum pernix g	gb-ba:AC004380	82.50	93.34	7.8e+03	120950	AC004380 Drosophila melano
gb-pat:AF121338	85.00	130.63	65.04	1511	AF121338 Sequence 1254 from Pa	gb-ba:DP09032	82.50	85.04	2.3e+04	120950	DP09032 Synchocystis sp.
gb-ov:AF128510	85.00	130.63	65.04	1511	AF128510 Sequence 1254 from Pa	gb-in:AC092217	82.50	82.57	3.1e+04	16347	AC092217 Drosophila melan
gb-ba:AE004885	85.00	110.59	849.65	13223	AE004885 Pseudomonas aeruginosa	gb-hg:AC010836	82.50	82.18	3.3e+04	164974	AC010836 Homo sapiens c
gb-ba:SM0591787	85.00	87.49	1.6e+04	149870	AC024579 Homo sapiens chromo	gb-in:AC092230	82.50	81.65	3.3e+04	165167	AC092230 Drosophila melan
gb-pat:AX127146	85.00	79.41	4.6e+04	349980	AX127146 Sequence 7062 from	gb-hg:AC018064	82.50	79.49	4.6e+04	218734	AC018064 Drosophila melan
gb-ba:NM00664K	84.50	124.57	141.46	2276	X77920 N.meningitidis outer me	gb-in:AE003577	82.50	77.87	5.8e+04	257662	AE003577 Drosophila melan
gb-ba:AE008779	84.50	103.58	2.1e+03	25184	AE008779 Salmonella typhimur	gb-hg:AC0108494	82.50	77.09	6.2e+04	281457	AC108494 Homo sapiens chr
gb-pr:NC156	84.50	95.43	6.3e+03	61883	AL153832 Neutropora citrissa DN	gb-ba:CN00000	82.50	76.35	6.8e+04	304290	AC098216 Rattus norvegicus
gb-pr:AL589734	84.50	94.03	6.8e+03	65882	AL589734 Human DNA sequence 4	gb-hg:AC098216	82.50	76.35	6.8e+04	304290	AC098216 Rattus norvegicus
gb-hg:AL592186	84.50	85.52	2.1e+04	168018	AL592186 Homo sapiens chromo	gb-pat:AX041921	82.50	75.02	8.1e+04	349980	AX041921 Sequence 816 fro
gb-hg:AC013703	84.50	85.04	2.3e+04	176704	AC013703 Homo sapiens clone	gb-com:AF117310	82.00	134.08	41.77	644	AF117310 Bos javanicus prion
gb-ba:AL646077	84.50	84.40	2.4e+04	189050	AL646077 Ralstonia solanacea	gb-com:AF117327	82.00	134.02	42.09	648	AF117327 Bos taurus prion
gb-com:OM074271	84.50	73.66	34.15	790	AF116805 Banna virus segment 1	gb-com:BOV00033	82.00	133.21	56.73	771	AF116805 Bovine prion
gb-ba:AE004815	84.00	120.80	229.49	10254	AF116805 Banna virus segment 1	gb-com:PI00000	82.00	132.37	52.29	771	AF116805 Bovine prion
gb-ba:AE004815	84.00	111.26	780.50	10254	AF116805 Banna virus segment 1	gb-pat:AX299246	82.00	131.97	54.77	804	AX299246 Sequence 3 from Pat
gb-ba:AE004815	84.00	105.92	1.5e+03	17971	AB033991 Bacillus circulans	gb-ro:RAT06268	82.00	129.11	79.06	1086	M15685 Rat 6-phospho
gb-hg:AC012890	84.00	90.01	1.2e+04	95614	AC012890 Drosophila melanog	gb-ro:RAT06268	82.00	124.92	135.35	1086	M15685 Rat 6-phospho
gb-in:AC008182	84.00	84.52	2.4e+04	170057	AC008182 Drosophila melanog	gb-ro:RAT06268	82.00	124.92	135.35	1086	M15685 Rat 6-phospho
gb-in:AE0008123	84.00	80.21	4.2e+04	267547	AE0008123 Drosophila melanog	gb-pr:AY058170	82.00	120.94	156.29	2561	AY058170 Arabidopsis thalia
gb-pat:ED08061	83.50	125.65	123.23	2061	E08061 DNA encoding variant cy	gb-pr:AY035014	82.00	120.71	232.09	2624	AY035014 Arabidopsis thalia
gb-pat:123834	83.50	125.65	123.23	2061	I23834 Sequence 9 from patent	gb-ba:BNP00000	82.00	120.71	232.09	2624	BNP00000 R. norvegicus mRNA fo
gb-pat:143345	83.50	125.65	123.23	2061	I43345 Sequence 5 from patent	gb-ba:BNP00000	82.00	120.71	232.09	2624	BNP00000 R. norvegicus mRNA fo

gb_ba:AE009873	-	82.00	106.66	1.4e+03	11454	AE009873	Pyrobaculum aerophilum	gb_pat:AN026013	80.50	123.92	153.83	1420	AX076013	Sequence 7 from
gb_pl:AC079829	+	82.00	86.38	1.9e+04	96699	AE009829	Arabidopsis thaliana	gb_vt:AE391158	80.50	122.54	183.94	1420	AF391158	Arabidopsis thaliana
gb_ba:DS0902	+	82.00	86.16	2.5e+04	122056	DS0902	Synechocystis sp. PCC	gb_vt:AV007235	80.50	122.50	184.63	1649	AV007235	Avian infectious b
gb_hg:AC019264	+	82.00	82.93	3.0e+04	138975	AC019264	Homo sapiens clone	gb_pat:AL15594	80.50	120.47	239.40	2040	AL15594	Homo sapiens clone
gb_hg:CN019R2	+	82.00	88.71	3.0e+04	142820	AL160235	Homo sapiens chromo	gb_pl:HPHYESA	80.50	120.16	249.17	2108	X86788	B. pluvialis mRNA for
gb_ln:AC0093192	+	82.00	81.19	3.7e+04	166857	AC093192	Drosophila melanog	gb_ba:BS020440	80.50	119.93	256.70	2160	FE220440	Bacillus megateries
gb_ba:AC104246	+	82.00	80.99	3.8e+04	170389	AC104246	Homo sapiens chromo	gb_ba:BS020462	80.50	114.15	358.32	3962	U02562	Bacillus subtilis N
gb_ln:AC010038	+	82.00	80.89	3.8e+04	172173	AC010038	Drosophila melanog	gb_ba:BA00CRX	80.50	113.68	571.68	4162	D45048	Bacillus subtilis g
gb_hg:AC104629	+	82.00	80.84	3.9e+04	173014	AC104629	Homo sapiens chromo	gb_pr:HS00AL48	80.50	107.91	1.2e+03	7637	X15943	Huamn calicivirus/Al
gb_hg:AC013632	+	82.00	80.72	3.9e+04	175250	AC013632	Homo sapiens clone	gb_ba:AE004348	80.50	104.19	1.9e+03	11288	AE004348	Vibrio cholerae
gb_ln:AC105264	+	82.00	80.60	4.0e+04	177581	AC105264	Drosophila melanog	gb_ba:AE008890	80.50	98.02	4.3e+03	21582	AE008890	Salmonella typhimur
gb_hg:AC034149	+	82.00	79.70	4.4e+04	195189	AC034149	Homo sapiens chromo	gb_hg:AL450383	80.50	83.89	2.6e+04	95271	AL450383	Homo sapiens chro
gb_ln:AE005557	+	82.00	77.05	6.2e+04	257867	AE005557	Drosophila melanog	gb_vt:NPPT3COMP	80.50	81.04	3.7e+04	128413	NPPT3COMP	Bombix mori nucle
gb_ln:AE003472	+	82.00	75.50	7.6e+04	303191	AE003472	Drosophila melanog	gb_pl:AP002485	80.50	80.49	4.0e+04	136150	AP002485	Oryza sativa ge
gb_rc:MMPSP1NGO	+	81.50	124.38	144.96	1627	AJ222800	Mus musculus mRNA for	gb_pr:CN05078N	80.50	79.00	4.9e+04	159171	AL591770	Human Chromosom
gb_rc:BC010978	+	81.50	120.36	245.39	1631	BC010978	Mus musculus, Simla	gb_hg:AC0091802	80.50	78.61	5.1e+04	165907	AC0093802	Homo sapiens ch
gb_ln:U60988	+	81.50	120.94	235.54	2337	U60988	Hyphantaria cunea storag	gb_ba:BS00B010	80.50	76.24	6.9e+04	212601	BS00212	Bacillus subtilis
gb_ba:AF177860	+	81.50	114.94	486.67	4388	AF177860	Agrobacterium tumefac	gb_ba:AL627279	80.50	74.18	9.0e+04	264050	AL627297	Salmonella enter
gb_ba:KPMRKA	+	81.50	110.72	836.14	6636	M55912	K. pneumoniae mrkA-EI	gb_ba:RPXK02	80.50	72.58	1.1e+05	312430	AJ235771	Rickettsia prow
gb_ba:AF232237	+	81.50	110.12	902.45	7277	AF232237	Agrobacterium tumefac	gb_hg:PEPMALJP3	80.50	72.41	1.1e+05	318221	AL049184	Pistomidium falc
gb_ba:AE009201	+	81.50	105.85	1.6e+03	11399	AE009201	Agrobacterium tumefa	gb_pl:OSA307662	80.00	71.78	1.2e+05	339972	AJ307662	Oryza sativa ge
gb_ba:AE008167	+	81.50	105.32	1.7e+03	12060	AE008167	Agrobacterium syringae	gb_rc:AF1139445	80.00	130.70	64.47	635	AF113945	Cervus elaphus can
gb_ba:AF030414	+	81.50	97.36	4.6e+03	27818	AF030414	Glucocorticoid acter di	gb_pat:AX057697	80.00	130.57	65.58	644	AX056297	Cervus elaphus n
gb_vt:SVY3434259	+	81.50	95.05	6.2e+03	35450	AF034259	Sulfobolus virus_ SIR	gb_rc:AP016328	80.00	128.85	81.70	771	AP016328	Cervus elaphus n
gb_hg:AC108989	+	81.50	93.74	7.4e+03	40684	AC108989	Rattus norvegicus c1	gb_rc:TSBR1R	80.00	128.85	81.70	771	X74759	T. streptococcus gene
gb_ln:CBG814F09	+	81.50	93.73	7.4e+03	40722	AC084501	Caenorhabditis briggs	gb_rc:TSBR1P	80.00	128.56	84.82	795	X74771	T. streptococcus gene
gb_ba:AF232004	+	81.50	91.32	1.0e+04	52498	AF232004	Pseudomonas syringae	gb_rc:TSBR1P	80.00	128.56	84.82	795	X74771	T. streptococcus gene
gb_hg:AC097203	+	81.50	82.08	3.3e+04	138561	AC097203	Rattus norvegicus c	gb_pat:AX2992248	80.00	128.45	85.99	800	AX2992248	Sequence 5 from
gb_hg:AC091291	+	81.50	81.03	3.7e+04	154728	AC097991	Rattus norvegicus c	gb_rc:AF156182	80.00	128.15	89.40	830	AF156182	Cervus elaphus n
gb_hg:AC0151203	-	81.50	80.24	4.1e+04	168004	AC0159203	Homo sapiens chromo	gb_rc:AF156187	80.00	127.88	92.56	854	AF156187	Articlocapra ametic
gb_hg:AC021645	-	81.50	79.93	4.3e+04	173590	AC051618	Homo sapiens chromo	gb_pat:AR142126	80.00	126.38	112.22	1000	AR151526	Sequence 25 from
gb_hg:AC021645	-	81.50	79.25	4.7e+04	183570	AC021645	Homo sapiens chromo	gb_vt:AL1029519	80.00	121.31	214.96	1703	U29519	Avian Infectious b
gb_ba:AL133486	+	81.50	79.10	4.8e+04	189370	AF010496	Rhodococcus capsula	gb_ba:SYDPD3	80.00	121.04	227.54	1784	X55289	Synechococcus pds g
gb_pr:AL133486	+	81.50	79.09	4.8e+04	189672	AF010496	Rhodococcus capsula	gb_ba:SYDPD3	80.00	121.04	227.54	1784	X55289	Synechococcus pds g
gb_pr:AC0792371	+	81.50	78.56	5.1e+04	200475	AC079271	Homo sapiens chromo	gb_pat:AX073186	80.00	120.34	243.49	1866	AX073186	Sequence 303 from
gb_rc:FCPR10N	-	81.00	137.52	26.87	373	V13698	Felis catus gene encodit	gb_rc:HAMPR	80.00	120.18	248.54	1818	M37381	Hamster scricpe pri
gb_rc:AF117328	+	81.00	133.00	46.01	600	AF117328	Bison bonasus prion pr	gb_rc:HAMPR	80.00	120.18	248.54	1818	K02234	Syrian golden hamst
gb_pat:AR87344	+	81.00	132.09	53.93	660	AR87344	Sequence 1 from Patent W	gb_rc:BOVPR1	80.00	119.93	276.98	2096	D10612	Bovine mRNA for pri
gb_pat:AR87344	+	81.00	132.09	53.93	660	AR87344	Sequence 1 from Patent W	gb_rc:BOVPR1	80.00	119.93	276.98	2096	D10612	Bovine mRNA for pri
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Rattus norvegicus
gb_rc:AF367625	+	81.00	131.26	59.98	720	AF367625	Microtus agrestis prid	gb_rc:AF406391	80.00	115.37	460.38	3178	AF406391	Ratt

gb_hlg:LMFCH36_25	79.50	80.76	3.9e+04	110000	Continuation (26 of 36) of 1	gb_ba:AE005504	78.50	100.00	3.3e+03	12115	AE005504 Escherichia coli
gb_pr:AC005383	79.50	79.69	4.4e+04	123110	AC005383 Homo sapiens chromc	gb_ba:AE007423	78.50	99.32	3.6e+03	13002	AE007423 Streptococcus pne
gb_hlg:AP004558	79.50	77.88	5.6e+04	148885	AP004558 Oryza sativa chromc	gb_ba:AE003997	78.50	98.82	3.8e+03	13714	AE003997 Xylella fastidiosa
gb_hlg:AP004665	79.50	77.29	6.0e+04	158299	AP004665 Oryza sativa chromc	gb_hlg:AC013288	78.50	94.13	6.8e+03	21972	AC013288 Diosphila melano
gb_hlg:AC091646	79.50	75.52	7.6e+04	190777	AC091646 Homo sapiens chromc	gb_ba:AE004669	78.50	94.13	7.0e+03	22444	AE004669 Pseudomonas aeru
gb_hlg:AC037476	79.50	75.17	7.9e+04	197872	AC037476 Homo sapiens chromc	gb_pr:AE1590293	78.50	83.59	2.7e+04	67911	AE1590293 Human DNA sequenc
gb_pr:AC017093	79.50	74.32	8.1e+04	203043	AC017093 Homo sapiens BAC cl	gb_pr:NCB101	78.50	82.51	3.1e+04	76702	AE1535927 Neurospora crassa
gb_ba:BSUB0016	79.50	74.35	8.1e+04	215640	AC017093 Bacillus subtilis com	gb_hlg:SPNU01910	78.50	82.51	3.2e+04	77743	AE1445932 Rattus norvegicus
gb_ba:AC008220	79.50	74.16	9.0e+04	220060	AC008220 Bacillus subtilis com	gb_hlg:AC091350	78.50	79.64	5.9e+04	102835	AC091350 Rattus norvegicus
gb_ba:FA001820	79.50	71.28	1.3e+05	297750	AL445066 Thermoplasma acidoph	gb_hlg:AC021947	78.50	77.51	4.5e+04	128574	AC021947 Homo sapiens clo
gb_ba:AP000006	79.50	70.62	1.4e+05	319000	AP000006 Pyrococcus horikosh	gb_hlg:AC044793	78.50	76.60	6.6e+04	141579	AC044793 Homo sapiens chr
gb_pat:AX044030	79.50	69.74	1.6e+05	349980	AX044030 Sequence 109 from H	gb_hlg:AC079981	78.50	76.27	6.9e+04	145166	AC079981 Homo sapiens chr
gb_com:AF113940	79.00	129.06	79.34	627	AF113940 Camelus dromedarius pr	gb_pr:AL391117	78.50	75.44	7.6e+04	159940	AL391117 Human DNA sequen
gb_com:AF009180	79.00	128.57	84.68	660	AF009180 Odocolleus hemionus he	gb_ba:AL591973	78.50	73.31	1.0e+05	200050	AL591973 Listeria monocyt
gb_com:AF117313	79.00	128.30	87.66	679	AF117313 Gazella subgutturosa R	gb_hlg:AC098730	78.50	71.99	1.2e+05	229788	AC098730 Mus musculus chr
gb_pr:FGU75383	79.00	127.83	93.21	714	U75383 Theropithecus gelada pri	gb_ba:AP000064	78.50	70.27	1.3e+05	247695	AP000064 Aeropyrum pernix
gb_pr:CMU75385	79.00	127.51	97.05	738	U75385 Cercopithecus torquatus aty	gb_ba:AP0002562	78.50	71.44	1.4e+05	270365	AP0002562 Escherichia coli
gb_com:AF009181	79.00	127.09	102.37	771	AF009181 Odocolleus hemionus he	gb_hlg:AE003432	78.50	68.85	1.8e+05	319551	AE003432 Drosophila melan
gb_com:AF117315	79.00	127.09	102.37	771	AF117315 Capra hircus cell-1 line	gb_com:AF117309	78.00	127.96	91.57	585	AF117309 Adax nasomaculatus
gb_com:AF195247	79.00	127.09	102.37	771	AF195247 Ovis aries prion prote	gb_com:AF117310	78.00	127.63	95.60	606	AF117310 Capra ibex nubiana
gb_pr:SSU08310	79.00	126.95	104.32	783	U08310 Salinili sciureus prion R	gb_com:OHU97331	78.00	126.82	106.10	670	U97331 Odocolleus hemionus h
gb_com:AF156184	79.00	126.39	112.01	830	AF156184 Odocolleus virginianus	gb_com:AF117323	78.00	126.67	108.06	739	AF117323 Hippotragus niger p
gb_com:AF156185	79.00	126.39	112.01	830	AF156185 Odocolleus virginianus	gb_com:AF156186	78.00	125.74	121.80	759	AF156186 Odocolleus virginia
gb_com:OAA100681	79.00	125.01	133.79	960	AD000681 Ovis aries prp gene, a	gb_pr:AFU15164	78.00	125.49	125.84	771	AB060288 Ovis aries musimon
gb_Da:AF311321	79.00	120.69	232.75	1511	AF311321 Sinorhizobium fredii	gb_com:AB060288	78.00	125.34	128.27	771	AB060288 Ovis aries musimon
gb_ro:AF373861	79.00	119.93	256.66	1637	AF373861 Mus musculus N-methyl	gb_com:AF117316	78.00	125.34	128.27	771	AF117316 Capra hircus tissue
gb_p1:MEFNBP	79.00	119.24	280.39	1760	Z28024 Zee mays gene for Ferre	gb_com:AF117320	78.00	125.34	128.27	771	AF117320 Ovis montanus pri
gb_hlg:LDU17042	79.00	110.87	820.08	4242	U17042 Leishmania donovani put	gb_com:AF117330	78.00	125.34	128.27	771	AF166334 Ovis canadensis pri
gb_sy:AF072539	79.00	110.68	840.56	4324	AF072539 Cloning vector pCMTA	gb_com:AF166334	78.00	125.34	128.27	771	AF166334 Ovis canadensis pri
gb_v1:PMU55002	79.00	110.68	840.56	4326	U55002 Panicum mosaic virus (H	gb_com:CHPR1PR	78.00	125.34	128.27	771	X74758 C.hircus gene for pri
gb_p1:MEPDKR2	79.00	108.36	1.1e+03	5319	M56856 Corn pryvate, orthophos	gb_com:HU023965	78.00	125.34	128.27	771	U25665 Odocolleus hemionus h
gb_v1:MEPDKR2	79.00	107.62	1.1e+03	5319	M56856 Corn pryvate, orthophos	gb_com:HU023965	78.00	125.34	128.27	771	U25665 Odocolleus hemionus h
gb_v1:AF441393	79.00	102.68	2.3e+03	110019	AF441393 Beas lefifolli virus,	gb_com:AX14161	78.00	125.34	128.27	771	AX141616 Capra hircus tissue
gb_ba:AE0041819	79.00	102.56	2.3e+03	11272	AE0041819 Mycobacterium bovis	gb_pat:AX141613	78.00	125.34	128.27	771	AX141613 Sequence 1 from Pa
gb_ba:AE00347	79.00	101.56	2.3e+03	112295	AE00347 Escherichia coli K12	gb_pat:AX141615	78.00	125.10	132.14	790	AX141615 Sequence 13 from Pa
gb_ba:AE005623	79.00	97.74	3.0e+03	12295	AE005623 Escherichia coli O15	gb_com:OAA7335	78.00	124.95	134.80	803	X79912 O. aries prp gene, exo
gb_ba:AE007027	79.00	90.32	4.6e+03	17506	AE007027 Mycobacterium tuber	gb_com:OAPR33	78.00	124.95	134.80	803	X79912 O. aries prp gene, exo
gb_ba:AE008829	79.00	95.93	5.6e+03	20371	AE008829 Salmonella typhimuri	gb_com:CHPR1PR	78.00	125.34	128.27	771	U25665 Odocolleus hemionus h
gb_hlg:AC020814	79.00	92.75	8.4e+03	28446	AC020814 Mus musculus clone R	gb_pat:XX292949	78.00	124.94	135.00	804	XX292949 Sequence 6 from Pat
gb_ba:AE00347	79.00	91.50	9.8e+03	32437	Z25586 Mycobacterium tuberclu	gb_com:CHPR1PR	78.00	124.94	135.00	804	XX292949 Sequence 6 from Pat
gb_ba:AE00347	79.00	90.32	1.1e+04	36734	AL049497 Streptomyces coelicc	gb_ro:AO73661	78.00	123.40	164.44	945	AY073661 Mus musculus olfact
gb_ba:AE00347	79.00	87.99	1.5e+04	46936	AL357143 Human DNA sequence f	gb_com:OAA7335	78.00	123.25	167.63	960	AY000739 Ovis aries prp gene
gb_ba:AE00347	79.00	85.59	2.1e+04	60373	AL355679 Human DNA sequence f	gb_ba:AF103968	78.00	123.10	170.84	975	AF103968 Streptomyces griseo
gb_hlg:AC055823	79.00	77.29	6.0e+04	144391	AC055823 Homo sapiens clone	gb_ba:AF117968	78.00	122.66	180.73	1021	AF117968 Pseudomonas resino
gb_pr:AL357894	79.00	76.62	6.6e+04	154847	AL357894 Human DNA sequence	gb_ba:AS0232654	78.00	122.31	189.19	1060	AF152654 Agricultural soil
gb_hlg:AC024720	79.00	76.56	6.6e+04	155834	AC024720 Homo sapiens chromc	gb_p1:YSCAF1A	78.00	116.76	385.49	1899	D63449 Yeast AT91 gene for
gb_pr:AL35580	79.00	76.41	6.7e+04	158338	AL35580 Homo sapiens chromc	gb_pat:E06817	78.00	116.39	404.16	1974	E06817 DNA encoding alcohol
gb_pr:AC022166	79.00	76.15	7.0e+04	162695	AL353580 Human DNA sequence	gb_pat:E08050	78.00	116.39	404.16	1974	E08050 cDNA encoding beer y
gb_hlg:AC093136	79.00	75.15	7.2e+04	180759	AC022166 Homo sapiens chromc	gb_pat:121407	78.00	116.39	404.16	1974	I21407 Sequence 16 from pat
gb_ba:AP002937	79.00	74.92	8.2e+04	185300	AP000063 Aeropyrum pernix ge	gb_pat:1733372	78.00	116.39	404.16	1974	I73372 Sequence 16 from pat
gb_hlg:AC016168	79.00	74.52	8.6e+04	193211	AC016168 Homo sapiens chromc	gb_pat:1733372	78.00	116.39	404.16	1974	I73372 Sequence 16 from pat
gb_hlg:AC092508	79.00	74.41	8.7e+04	195411	AC092508 Mus musculus chromc	gb_com:SHPR3A	78.00	115.38	459.80	2194	D38179 Sheep gene for prion
gb_pr:AL591074	79.00	74.04	9.1e+04	203076	AL591074 Human DNA sequence	gb_com:SA6825	78.00	114.35	525.07	2446	SA6825 prion protein [mank,
gb_ba:AL627269	79.00	71.91	1.2e+05	254050	AL627269 Homo sapiens chromc	gb_com:OAPR1ON	78.00	109.62	962.99	4020	AJ323072 Ovis aries prp gen
gb_ba:AP002567	79.00	69.75	1.6e+05	318703	AP002567 Escherichia coli O1	gb_p1:AE062171	78.00	105.33	1.7e+05	6309	AE062171 Neurospora crassa
gb_ba:AP002937	79.00	69.43	1.6e+05	329709	AP002937 Mesorhizobium loti	gb_Da:DB7026	78.00	101.54	2.7e+03	9390	DB7026 Bacillus stearotherm
gb_ba:AP003012	79.00	68.86	1.7e+05	346547	AP003012 Mesorhizobium loti	gb_ba:AE009124	78.00	99.59	3.5e+03	11533	AE009124 Agrobacterium tum
gb_ov:AF0318161	78.50	68.88	1.7e+05	349498	AF038161 Ictailurus punctatus cl	gb_ba:AE000496	78.00	99.27	3.7e+03	11929	AE000808 Agrobacterium tum
gb_com:AF113943	78.50	124.29	40.67	330	AF038161 Ictailurus punctatus cl	gb_ba:AE000808	78.00	98.60	4.0e+03	12974	AE000808 Agrobacterium sp.
gb_pr:AF076976	78.50	126.33	112.96	762	AF076976 Homo sapiens prion prote	gb_com:OAPR1	78.00	90.05	1.2e+04	31412	U67922 Ovis aries prion pr
gb_com:CDPRCELF	78.50	126.25	114.05	768	AF076976 Homo sapiens prion prote	gb_hlg:AC012904	78.00	84.48	2.4e+04	56423	AC012904 Arabidopsis thaliana
em_hlg9_Inv:AC029300	78.50	123.50	162.42	1026	AC029300 Giardia intestinalis	gb_pr:AL352631	78.00	81.03	3.7e+04	81035	AL352631 Arabidopsis thaliana
em_hlg9_Inv:AC005487	78.50	123.52	170.18	1066	AC005487 Giardia intestinalis	gb_hlg:AC096097	78.00	79.44	4.6e+04	96474	AC096097 Human DNA sequenc
gb_pat:AF078994	78.50	122.52	184.12	1137	AF078994 Sequence 5 from Paten	gb_pr:AL354993	78.00	78.41	5.2e+04	106736	AL354993 Human DNA sequen
gb_ba:CDPOXINA	78.50	117.39	355.26	1948	AL17194 Clostridium difficile c	gb_hlg:AC093316_0	78.00	78.12	5.4e+04	110000	AC093316 Mus musculus chr
gb_pat:AL31365	78.50	117.39	355.26	1948	AL17194 Clostridium difficile c	gb_hlg:LMFCH331_12	78.00	78.12	5.4e+04	110000	Continuation (13 of 22) o
gb_ba:STRU0420A	78.50	114.75	498.50	2571	M80215 Streptococcus pneumonia	gb_hlg:LMFCH331_13	78.00	78.12	5.4e+04	110000	Continuation (14 of 22) o
gb_pat:BD003822	78.50	108.38	1.1e+03	5020	BD003822 Polynucleotide of Str	gb_hlg:LMFCH332_19	78.00	78.12	5.4e+04	110000	Continuation (20 of 28) o
gb_ba:BCY10545	78.50	108.37	1.1e+03	5025	Y10545 Cloning vector pZero-2T	gb_pr:AC0053301	78.00	77.81	5.7e+04	113688	AC0053301 Homo sapiens chr
gb_ba:AE008484	78.50	101.80	2.6e+03	10029	AE008484 Streptococcus pneum	gb_hlg:AC016515	78.00	77.23	6.1e+04	120851	AC016515 Homo sapiens chr
gb_ba:AE000890	78.50	100.82	3.0e+03	11107	AE000890 Methanobacterium the	gb_pr:AL157772	78.00	77.14	6.2e+04	121949	AL157772 Human DNA sequen

gb-pr:AL136379	-	78.00	76.69	6.5e+04	127816	! AL136379 Human DNA sequence	gb-ba:AP003011	+	77.50	66.32	2.4e+05	346510	! AP003011 Mesorhizobium 10
gb-pr:AC009655	-	78.00	76.17	7.0e+04	135033	! AC009655 Homo sapiens BAC cl	gb-pat:AX120085	+	77.50	66.22	2.4e+05	349980	! AX120085 Sequence 18 from Pa
gb-hg:AC009673	+	78.00	75.73	7.4e+04	141545	! AC009673 Trypanosoma brucei 1	gb-pat:AX120695	+	77.00	125.34	1.2e+08	641	! AX057695 Sequence 1 from Pa
gb-pr:AC009848	+	78.00	75.51	7.6e+04	144724	! AC009848 Oryza sativa chrom	gb-com:AB060290	+	77.00	123.58	160.72	771	! AB060290 Budorcas taxicolor
gb-pr:AC009810	+	78.00	75.37	7.7e+04	146671	! AC009810 Homo sapiens chrom	gb-pat:AX16167	+	77.00	123.58	160.72	771	! AX16167 Sequence 15 from Pa
gb-hg:AC008407	-	78.00	74.86	8.2e+04	155007	! AC008407 Trypanosoma brucei 1	gb-pat:AX299244	+	77.00	123.15	169.16	804	! AX299244 Sequence 1 from Pat
gb-hg:AL513545	-	78.00	74.66	8.4e+04	158297	! AL513545 Homo sapiens chrom	gb-com:OAA00680	+	77.00	121.49	210.04	960	! AU000679 Ovis aries PEP gene
gb-hg:AC0010565	-	78.00	74.56	8.5e+04	159928	! AC0010565 Homo sapiens melanog	gb-pat:AX137063	+	77.00	121.49	210.04	960	! AU000680 Ovis aries PEP gene
gb-hg:AC024009	+	78.00	73.29	1.0e+05	182662	! AC024009 Homo sapiens chrom	gb-pat:AX145127	+	77.00	118.45	310.42	1332	! AR145127 Sequence 27 from p
gb-hg:AC004091	+	78.00	73.08	1.0e+05	186883	! AC004091 Homo sapiens chrom	gb-pi:AY033947	+	77.00	117.95	331.47	1395	! AY033947 Arabidopsis thaliana
gb-hg:AC006999	+	78.00	72.69	1.1e+05	194713	! AC006999 Mus musculus chrom	gb-ba:KPMR86	+	77.00	117.85	335.24	1408	! X03146 Klebsiella pneumoniae
gb-hg:AC006908	+	78.00	72.52	1.1e+05	198192	! AC006908 Mus musculus chrom	gb-ba:AY15114	+	77.00	115.16	473.10	1867	! Y15114 Anabaena PCC7210 crt
gb-pr:AC021020	-	78.00	72.30	1.1e+05	202771	! AC021020 Homo sapiens BAC cl	gb-ro:MSRPR	+	77.00	114.05	545.19	2097	! M13685 Mouse p10n protein
gb-hg:AC0104135	-	78.00	71.53	1.3e+05	219940	! AC0104135 Homo sapiens clone	gb-ro:BC020074	+	77.00	113.87	558.23	2138	! BC020074 Mus musculus, Slm1
gb-hg:AC0104135	-	78.00	71.53	1.3e+05	220107	! AC0104135 Homo sapiens clone	gb-ro:AB041585	+	77.00	113.78	564.61	2158	! AB041585 Mus musculus brain
gb-hg:AL606909	+	78.00	71.35	1.3e+05	224061	! AL606909 Mus musculus chrom	gb-ba:AF051693	+	77.00	111.58	748.62	2719	! AF051693 Pseudomonas aerugi
gb-ba:AL627425	+	78.00	70.97	1.3e+05	227013	! AL627425 Mus musculus chrom	gb-in:COB278949	+	77.00	110.82	825.61	2946	! AJ278949 Cystosporogones op
gb-ba:AL627425	+	78.00	70.97	1.3e+05	233050	! AL627425 Salmonella enterica	gb-pat:AX123472	+	77.00	107.81	1.2e+03	4032	! AX123472 Sequence 3388 from
gb-ba:AP000995	+	78.00	69.26	1.7e+05	279160	! AP000995 Thermoplasma volcan	gb-ro:AB021131	+	77.00	106.17	1.2e+03	4800	! AB021131 Mus musculus RIM2
gb-ba:AP000345	+	78.00	69.16	1.7e+05	281893	! AP000345 Drosophila melanog	gb-pat:AX137063	+	77.00	105.82	1.6e+03	4980	! AX137063 Sequence 2 from Pa
gb-ba:AP002994	+	78.00	67.17	2.2e+05	347660	! AP002994 Mesorhizobium loti	gb-ro:DB6948	+	77.00	103.34	2.2e+03	6466	! DB6948 Mus musculus mRNA fo
gb-ba:AB074293	+	77.50	124.99	134.09	729	! AB074293 Photobacterium damela	gb-ba:AY008284	+	77.00	100.58	3.1e+03	8637	! AY008284 Yersinia enterocol
gb-pat:AX120221	+	77.50	122.73	179.09	924	! AX120221 Sequence 137 from Pat	gb-ba:AY008284	+	77.00	98.61	3.9e+03	10619	! AE002346 Neisseria meningi
em-hgco:Inv:AC081030	+	77.50	122.50	184.55	947	! AC081030 Giardia intestinalis c	gb-ba:AE008332	+	77.00	98.56	4.0e+03	10684	! AE008332 Agrobacterium tum
gb-pat:AX065883	+	77.50	121.79	202.06	1020	! AX065883 Sequence 227 from Pat	gb-ba:AE008778	+	77.00	91.51	9.8e+03	22388	! AE008778 Salmonella typhim
em-hgco:Inv:AC038755	+	77.50	121.47	210.55	1055	! AC038755 Giardia intestinalis	gb-hg:AC102930	+	77.00	83.73	2.7e+04	50734	! AC102930 Mus musculus c10n
gb-ba:BA000000	+	77.50	113.25	604.22	2502	! BA000000 Homo sapiens	gb-hg:AC102930	+	77.00	81.11	3.7e+04	56821	! AC102930 Mus musculus c10n
gb-ba:BA000000	+	77.50	112.33	680.22	2757	! BA000000 Homo sapiens	gb-pr:AC007722	+	77.00	79.92	4.3e+04	75703	! AC007722 Homo sapiens chro
gb-ba:BA000000	+	77.50	111.84	723.84	2901	! BA000000 Homo sapiens	gb-hg:AC007722	+	77.00	79.92	4.3e+04	75703	! AC007722 Homo sapiens chro
gb-pat:AL103929	+	77.50	107.75	1.2e+03	4460	! AL103929 Sequence 1 from Pat	gb-pr:AC007722	+	77.00	79.92	4.3e+04	75703	! AC007722 Homo sapiens chro
gb-ba:AE229829	+	77.50	107.69	1.2e+03	4487	! AE229829 Pseudomonas fluores	gb-pr:AL139277	+	77.00	72.88	1.1e+05	158591	! AL139277 Homo sapiens chro
gb-ba:AE229829	+	77.50	107.40	1.2e+03	4625	! AE229829 Pseudomonas fluores	gb-pr:AL139277	+	77.00	72.88	1.1e+05	158591	! AL139277 Homo sapiens chro
gb-ba:NE047945	+	77.50	106.00	1.5e+03	5559	! NE047945 N. crassa plasma membra	gb-ba:SY05SLH	+	77.00	74.62	8.5e+04	133106	! SY05SLH Drosophila melan
gb-ba:DB09277	+	77.50	105.44	1.6e+03	5684	! DB09277 Streptococcus mutans ge	gb-ba:SY05SLH	+	77.00	74.62	8.5e+04	133106	! SY05SLH Drosophila melan
gb-ba:STR0778	+	77.50	100.04	3.3e+03	10029	! STR0778 Streptococcus mutans ge	gb-pr:AC003043	+	77.00	74.10	9.0e+04	133488	! AC003043 Homo sapiens chr
gb-ba:AE007786	+	77.50	99.21	4.9e+03	13949	! AE007786 Clostridium acetobu	gb-pr:AL590032	+	77.00	72.90	1.0e+05	156060	! AL590032 Homo sapiens chr
gb-ba:AE005936	+	77.50	96.45	5.2e+03	14613	! AE005936 Caulobacter crescent	gb-hg:AC004153	+	77.00	72.90	1.0e+05	156060	! AC004153 Plasmodium falci
gb-v1:NEPUCEDCS	+	77.50	95.49	5.2e+03	16168	! NEPUCEDCS Drosophila melanog	gb-hg:AC004153	+	77.00	72.90	1.0e+05	156060	! AC004153 Plasmodium falci
gb-ba:AE004856	+	77.50	91.73	9.5e+03	24000	! AE004856 Pseudomonas aerugin	gb-hg:AC004153	+	77.00	72.90	1.0e+05	156060	! AC004153 Plasmodium falci
gb-ba:SC0778	+	77.50	86.22	1.1e+04	27437	! SC0778 Streptococcus coelic	gb-pr:AL139277	+	77.00	72.88	1.1e+05	158591	! AL139277 Homo sapiens chro
gb-hg:AC01247	+	77.50	80.46	1.9e+04	42845	! AC01247 Drosophila melanog	gb-hg:AC0137058	+	77.00	72.74	1.1e+05	163093	! AC0137058 Homo sapiens chr
gb-hg:AC010253_3	+	77.50	80.38	4.1e+04	9111	! AC010253_3 Continuation (4 of 4) of AC01	gb-in:AC011907	+	77.00	72.43	1.1e+05	164595	! AC011907 Drosophila melan
gb-hg:AC017776	+	77.50	78.45	5.2e+04	96887	! AC017776 Drosophila melanog	gb-hg:AC014423	+	77.00	72.47	1.1e+05	165622	! AC014423 Homo sapiens chr
gb-pr:AP003268	+	77.50	78.45	5.2e+04	96889	! AP003268 Oryza sativa genomi	gb-hg:AL160171	+	77.00	72.47	1.1e+05	165607	! AL160171 Homo sapiens chr
gb-hg:AC003985	+	77.50	78.34	5.3e+04	98002	! AC003985 Oryza sativa chrom	gb-pr:AC025094	+	77.00	72.38	1.1e+05	167218	! AC025094 Homo sapiens chr
gb-pr:AC003881	+	77.50	78.03	5.5e+04	101303	! AC003881 Homo sapiens chrom	gb-hg:AC025094	+	77.00	72.30	1.1e+05	168481	! AC025094 Homo sapiens chr
gb-pr:AP004511	+	77.50	77.33	6.0e+04	108987	! AP004511 Lotus japonicus gen	gb-hg:AC019758	+	77.00	72.19	1.2e+05	170575	! AC019758 Drosophila melan
gb-pat:AA8542	+	77.50	75.47	7.6e+04	132470	! AA8542 Sequence 1 from Pat	gb-hg:AC019758	+	77.00	71.80	1.2e+05	170575	! AC019758 Drosophila melan
gb-v1:L22858	+	77.50	75.37	7.7e+04	133894	! L22858 Autographa californic	gb-hg:AC019758	+	77.00	71.80	1.2e+05	170575	! AC019758 Drosophila melan
gb-hg:AC004754	+	77.50	74.98	8.1e+04	139434	! AC004754 Oryza sativa chrom	gb-hg:AC021540	+	77.00	71.65	1.2e+05	180520	! AC021540 Homo sapiens chr
gb-hg:AC021743	+	77.50	72.81	9.0e+04	152831	! AC021743 Homo sapiens chrom	gb-pr:AC069077	+	77.00	71.52	1.3e+05	183087	! AC069077 Homo sapiens chr
gb-hg:AC007189	+	77.50	72.82	1.1e+05	157584	! AC007189 Homo sapiens chrom	gb-hg:AC019151	+	77.00	71.19	1.3e+05	183985	! AC019151 Homo sapiens chr
gb-pr:AL451103	+	77.50	72.80	1.1e+05	175384	! AL451103 Human DNA sequence	gb-hg:AC019151	+	77.00	71.19	1.3e+05	183985	! AC019151 Homo sapiens chr
gb-in:AL510025	+	77.50	72.79	1.1e+05	175679	! AL510025 Drosophila melanog	gb-pr:CN50681	+	77.00	70.91	1.4e+05	195907	! CN50681 Homo sapiens chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus musculus chr
gb-hg:AC0100774	+	77.50	72.68	1.1e+05	177590	! AC0100774 Drosophila melanog	gb-hg:AC0193858	+	77.00	70.67	1.4e+05	200056	! AC0193858 Mus

gb_sts:HSHP076	-	76.50	131.88	55.45	294	AL158939 H. sapiens STS from fld	gb_hcg:AL645966	-	76.50	71.19	1.3e+05	172746	AL645966 Mus musculus chr
gb_pr:GGU08300	+	76.50	122.81	177.35	762	U08300 Gorilla gorilla piron pr	gb_pr:AA18272	-	76.50	70.85	1.4e+05	178952	AF418272 Homo sapiens coa
gb_pr:GGU15166	+	76.50	122.81	177.35	762	U08300 Gorilla gorilla piron pr	gb_hcg:AC012513	+	76.50	70.84	1.4e+05	179149	AF418272 Homo sapiens chr
gb_com:OAAJ738	+	76.50	122.47	185.33	790	AJ000738 Ovis aries PEP gene	gb_hcg:AC026331	+	76.50	70.84	1.4e+05	179172	AC026331 Homo sapiens chr
gb_vt:RVU22480	+	76.50	117.37	356.48	1350	U22480 Rabies virus 91350MA nu	gb_hcg:AC027290	+	76.50	70.78	1.4e+05	180339	AC027290 Homo sapiens chr
gb_vt:RVU22481	+	76.50	117.37	356.48	1350	U22480 Rabies virus 91350MA nu	gb_pr:AP002797	+	76.50	70.52	1.4e+05	185319	AP002797 Homo sapiens chr
gb_vt:RVU22633	+	76.50	117.37	356.48	1350	U22633 Rabies virus 8721AFA nu	gb_hcg:AC094898	+	76.50	70.41	1.4e+05	187368	AC094898 Rattus norvegicus chr
gb_vt:AF167478	+	76.50	117.16	356.85	1350	U22633 Rabies virus 8721AFA nu	gb_hcg:AL665807	+	76.50	70.30	1.5e+05	189566	AL665807 Rattus norvegicus chr
gb_pr:AY063633	+	76.50	115.38	459.82	1663	U4063833 Arabidopsis thaliana	gb_hcg:AC025558	+	76.50	70.29	1.5e+05	189748	AC025558 Homo sapiens chr
gb_pr:HMMR9329	+	76.50	115.35	461.84	1663	U4063833 Arabidopsis thaliana	gb_hcg:AL607030	+	76.50	70.29	1.5e+05	189748	AL607030 Mus musculus chr
gb_ba:SNYNEUR	+	76.50	114.61	507.50	1803	M55342 S. typhimurium neuramin	gb_hcg:AC021155	+	76.50	69.93	1.5e+05	197031	AC021155 Homo sapiens chr
gb_ro:MM082441	+	76.50	114.40	521.62	1844	U82441 Mus musculus EBR-7 mRNA	gb_pr:BSB00002	+	76.50	68.92	1.5e+05	197466	BSB00002 Homo sapiens chr
gb_pr:AF428342	+	76.50	113.99	549.73	1925	AF428342 Arabidopsis thaliana	gb_pr:AL539930	+	76.50	68.92	1.5e+05	201129	AL539930 Homo sapiens chr
gb_pr:AF428352	+	76.50	113.64	575.28	1958	AF37352 Arabidopsis thaliana	gb_hcg:AL645807	+	76.50	69.55	1.6e+05	205077	AL645807 Mus musculus chr
gb_ro:AF42480	+	76.50	112.94	629.15	2150	AF142480 Rattus norvegicus rec	gb_hcg:AC067917	+	76.50	69.23	1.7e+05	221560	AC067917 Homo sapiens chr
gb_ba:BCACMLAY	+	76.50	112.32	680.97	2294	K00563 b. subtilis alpha-amyl	gb_ba:BSUB0020	+	76.50	69.23	1.7e+05	221560	BSUB0020 Bacillus subtilis
gb_pr:HSR0930	+	76.50	111.17	789.71	2590	U80930 Human stem cell fam	gb_ba:BSUB0002	+	76.50	68.84	1.8e+05	221560	BSUB0002 Bacillus subtilis
gb_pr:STNPLMYL	+	76.50	111.03	803.88	2528	K00562 b. subtilis alpha-amyl	gb_pr:AL539930	+	76.50	68.92	1.5e+05	226860	AL539930 Human DNA sequen
gb_ro:MM071189	+	76.50	110.97	809.48	2643	U71189 Mus musculus met-mesenc	gb_hcg:AC096350	+	76.50	68.22	1.9e+05	235932	AC096350 Rattus norvegicus
gb_ba:AF105339	+	76.50	110.94	812.85	2652	AF105339 Neisseria meningitidis	gb_hcg:AC020885	+	76.50	67.00	2.2e+05	268294	AC020885 Mus musculus clo
gb_ba:BSAMYL	+	76.50	110.84	833.34	2680	V00101 Bacillus subtilis amy	gb_hcg:AC073671	+	76.50	66.69	2.5e+05	276938	AC073671 Mus musculus act
gb_ba:AF012211	+	76.50	109.83	937.20	2980	AF012211 Neisseria meningitidis	gb_ba:RAC104	+	76.50	66.00	2.5e+05	297750	AL445066 Thermoplasma act
gb_ro:AA116519S1	+	76.50	109.58	967.23	3058	AF116519 Mus musculus periplak	gb_ba:AP003129	+	76.50	66.00	2.5e+05	298050	AP003129 Staphylococcus a
gb_pr:KPLPER92	+	76.50	108.82	1.1e+03	3315	U98501 Cloning vector pATL PCR	gb_hcg:AC078979	+	76.50	63.29	2.7e+05	321128	AC078979 Homo sapiens chr
gb_ba:PM18358	+	76.50	108.81	1.1e+03	3318	U18528 Neisseria meningitidis	gb_ba:AP003138	+	76.50	63.29	2.7e+05	321128	AP003138 Staphylococcus a
gb_pr:185759	+	76.50	108.81	1.1e+03	3318	U18528 Neisseria meningitidis	gb_ba:AP003138	+	76.50	63.29	2.7e+05	321128	AP003138 Staphylococcus a
gb_ba:AM022174	+	76.50	108.30	3499	AB022174 Aeromonas caviae gene	U927203 Mus musculus Olf-1/EBF-	gb_com:AF090852	+	76.50	64.47	3.0e+05	349999	AF090852 Klebsiella oxytoca
gb_ro:MM092703	+	76.50	107.60	1.2e+03	3768	U927203 Mus musculus Olf-1/EBF-	gb_com:AF090852	+	76.50	64.47	3.0e+05	349999	AF090852 Klebsiella oxytoca
gb_ba:RSMGPN	+	76.50	107.27	1.3e+03	3900	U250182 R. sphaeroides mgs gene	gb_com:AF117312	+	76.50	123.00	173.07	681	AF117312 Equus caballus tist
gb_in:DDU72236	+	76.50	106.02	1.5e+03	4446	U72236 Dictyostellium mgs gene	gb_com:AF117312	+	76.50	123.00	173.07	681	AF117312 Equus caballus tist
gb_pat:AF139771	+	76.50	103.39	2.1e+03	5864	AF139771 Sequence 47 from pat	gb_com:AF117318	+	76.50	123.00	173.07	681	AF117318 Equus caballus cell
gb_pat:AF140290	+	76.50	103.39	2.1e+03	5864	AF140290 Sequence 47 from pat	gb_pr:CAW75384	+	76.50	122.55	183.36	714	U75384 Cercopithecus aethiops
gb_ba:BCSCPOFA	+	76.50	102.68	2.3e+03	5864	AF140290 Sequence 47 from pat	gb_pr:MSU75382	+	76.50	122.55	183.36	714	U75384 Cercopithecus aethiops
gb_ba:CVU46197	+	76.50	101.73	2.6e+03	6882	M20309 Bacillus subtilis spo	gb_pr:AM088312	+	76.50	122.43	186.18	723	U08312 Callicebus molloch pri
gb_sy:CVU46197	+	76.50	100.92	2.6e+03	6882	U46197 Cloning vector pDG162	gb_pr:MSU08303	+	76.50	122.43	186.18	723	U08312 Callicebus molloch pri
gb_ba:AE006124	+	76.50	98.28	4.1e+03	10029	U4606724 Sulfolobus solfatar	gb_pr:CAU08291	+	76.50	122.24	190.91	738	U08291 Cercopithecus aethiops
gb_ba:AE008295	+	76.50	98.16	4.2e+03	10156	U46196 Cloning vector pDG162	gb_pr:CGU75386	+	76.50	122.24	190.91	738	U08291 Cercopithecus aethiops
gb_sy:CVU46196	+	76.50	97.81	4.3e+03	10426	AX100838 Sequence 82 from pat	gb_pr:CGU75386	+	76.50	122.24	190.91	738	U08291 Cercopithecus aethiops
gb_pat:AX100838	+	76.50	97.81	4.3e+03	10426	AX100838 Sequence 82 from pat	gb_pr:CGU75386	+	76.50	122.24	190.91	738	U08291 Cercopithecus aethiops
gb_ba:AE004629	+	76.50	97.60	4.5e+03	10776	U46198 Cloning vector pDG1726	gb_pr:CGU75386	+	76.50	122.24	190.91	738	U08291 Cercopithecus aethiops
gb_sy:CVU46198	+	76.50	97.60	4.5e+03	10776	U46198 Cloning vector pDG1726	gb_pr:SSU15165	+	76.50	121.97	197.56	759	U15165 Saimiri sciureus majo
gb_ba:AE009319	+	76.50	96.97	4.9e+03	11355	AE009319 Agrobacterium tumef	gb_pr:CGU08297	+	76.50	121.93	198.52	762	U08297 Colobus guereza pri
gb_in:CEH19M07	+	76.50	96.60	5.1e+03	11966	U292833 Caenorhabditis elegans	gb_pr:MFU08302	+	76.50	121.93	198.52	762	U08297 Colobus guereza pri
gb_ba:RLHUPGENE	+	76.50	92.83	8.3e+03	17786	X55974 R. leguminosarum bv. v	gb_pr:MFU08302	+	76.50	121.93	198.52	762	U08297 Colobus guereza pri
gb_pr:AL391420	+	76.50	92.60	8.5e+03	18216	AL391420 Human DNA sequence	gb_pr:MFU08302	+	76.50	121.93	198.52	762	U08297 Colobus guereza pri
gb_sy:AE246703	+	76.50	91.08	1.0e+04	21376	AE246703 Synthetic construct	gb_pr:MFU08302	+	76.50	121.93	198.52	762	U08297 Colobus guereza pri
gb_ba:AE008739	+	76.50	90.56	1.1e+04	22578	AE008739 Salmonella typhimur	gb_pr:MFU08302	+	76.50	121.93	198.52	762	U08297 Colobus guereza pri
gb_ba:BSMNA320D	+	76.50	89.23	1.3e+04	25940	Z45782 B. subtilis chromosome	gb_pr:MFU08302	+	76.50	121.93	198.52	762	U08297 Colobus guereza pri
gb_in:CHBCG41N04	+	76.50	86.31	1.9e+04	34978	AC084590 Caenorhabditis brig	gb_ro:AF117322	+	76.50	121.89	199.47	765	AF117322 Rattus norvegicus p
gb_ba:AOCCZAK361	+	76.50	84.73	2.3e+04	41625	ALU49826 Streptomyces coelic	gb_com:MFU08852	+	76.50	121.78	202.34	774	U08522 Mus musculus p
gb_hcg:AC102954	+	76.50	81.12	3.7e+04	60864	AC102954 Rattus norvegicus ch	gb_pat:AX299245	+	76.50	121.42	231.99	804	AX299245 Sequence 2 from pat
gb_hcg:AC106124	+	76.50	79.97	4.9e+04	71497	AC106124 Rattus norvegicus c1	gb_pr:CLM07PB	+	76.50	120.47	257.17	888	CLM07PB Rattus norvegicus
gb_hcg:AL606703	+	76.50	77.97	5.5e+04	84711	AP006703 Dantio rerio chromos	gb_ro:AY074028	+	76.50	119.91	257.17	942	AY074028 Mus musculus olfact
gb_pl:AP002067	+	76.50	76.42	6.7e+04	98653	AC103038 Rattus norvegicus c1	gb_ba:CCN28334	+	76.50	118.87	258.50	946	CCN28334 Oryzopsis sp. partia
gb_hcg:AC103038	+	76.50	76.42	6.7e+04	98653	AC103038 Rattus norvegicus c1	gb_ba:SSR288128	+	76.50	118.87	258.50	946	SSR288128 Rattus norvegicus
gb_pr:AC005006	+	76.50	76.01	7.1e+04	101170	AC005006 Homo sapiens clone	gb_com:OAAJ736	+	76.50	117.54	348.76	1209	AX135431 Sequence 9 from pa
gb_hcg:AC022426	+	76.50	75.43	7.6e+04	110559	AC011469 Homo sapiens chrom	gb_ba:CLM07PB	+	76.50	117.34	357.94	1235	AY07100 Clonitridium thermace
gb_pr:AC011629	+	76.50	74.94	8.1e+04	116458	AC106626 Rattus norvegicus c	gb_ro:AB047002	+	76.50	116.24	412.06	1386	J02720 Rattus norvegicus m
gb_hcg:AC106226	+	76.50	74.74	8.2e+04	117735	AC068274 Homo sapiens BAC c1	gb_pr:RATARB1	+	76.50	116.24	412.06	1386	J02720 Rattus norvegicus m
gb_pr:AC0068274	+	76.50	72.78	1.1e+05	118955	D50453 Bacillus subtilis DNA	gb_pr:AF435457	+	76.50	115.15	528.34	1699	AF435457 Brachycephalus
gb_ba:DS0453	+	76.50	72.78	1.1e+05	118955	D50453 Bacillus subtilis DNA	gb_pr:AF435457	+	76.50	115.15	528.34	1699	AF435457 Brachycephalus
gb_hcg:AC109808	+	76.50	72.46	1.1e+05	150154	AC109808 Homo sapiens chrom	gb_pr:AF435457	+	76.50	115.15	528.34	1699	AF435457 Brachycephalus
gb_hcg:AP003217	+	76.50	72.46	1.1e+05	151677	AC003217 Oryza sativa chrom	gb_pr:AF435457	+	76.50	115.15	528.34	1699	AF435457 Brachycephalus
gb_pr:AC023118	+	76.50	72.26	1.2e+05	154339	AC023118 Homo sapiens BAC c1	gb_vt:AY034455	+	76.50	109.67	956.58	2763	AY034455 Equine herpesvirus
gb_hcg:AC006513	+	76.50	71.88	1.2e+05	160574	AC006513 Homo sapiens chrom	gb_pat:AX037558	+	76.50	109.33	999.87	2865	AX037558 Sequence 13 from p
gb_pr:AC079965	+	76.50	71.79	1.2e+05	163494	AC079965 Homo sapiens 12q B	gb_pr:AX037558	+	76.50	108.61	1.1e+03	3090	D50093 Rattus norvegicus DN
gb_hcg:AC031978	+	76.50	71.41	1.3e+05	167359	AC031978 Homo sapiens chrom	gb_pat:AX037559	+	76.50	108.56	1.1e+03	3104	AX037559 Sequence 13 from p


```

gb_ba:AF065404      +      75.50      68.95      1.7e+05      181654      ! AF065404 Bacillus anthracis
gb_pr:AC079239      +      75.50      68.91      1.7e+05      182406      ! AC079239 Homo sapiens BAC C1
gb_in:AC093198      +      75.50      68.88      1.7e+05      182901      ! AC093198 Drosophila melanoga
gb_ba:AP000063      -      75.50      68.76      1.8e+05      185300      ! AP000063 Aeropyrum pernix ge
gb_hng:AC079560      +      75.50      68.60      1.8e+05      188406      ! AC079560 Mus musculus clone
gb_pr:AC022113      -      75.50      68.59      1.8e+05      188570      ! AC022113 Homo sapiens chromc
gb_hng:AC073658      +      75.50      68.58      1.8e+05      188873      ! AC073658 Mus musculus clone
gb_in:AC093440      +      75.50      68.43      1.9e+05      192132      ! AC093440 Drosophila melanoga
gb_hng:AC025784      -      75.50      68.33      1.9e+05      193961      ! AC025784 Mus musculus clone
gb_pr:AL358975      +      75.50      68.21      1.9e+05      196337      ! AL358975 Human DNA sequence
gb_hng:AC013569      +      75.50      68.12      1.9e+05      198160      ! AC013569 Homo sapiens clone

```

seq_name: gb_ba:AF242417

seq_documentation_block:

```

LOCUS      AF242417              777 bp      DNA      linear      BCT 01-JAN-2002
DEFINITION Escherichia coli heat-labile enterotoxin subunit A precursor, gene,
complete cds.
ACCESSION  AF242417
VERSION    AF242417.1 GI:18026885
KEYWORDS
SOURCE
ORGANISM   Escherichia coli.
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.

```

REFERENCE 1 (bases 1 to 777)

AUTHORS Chang, G.N. and Ho, K.C.

TITLE Heat-labile enterotoxin subunit A gene of Escherichia coli

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 777)

AUTHORS Chang, G.N. and Ho, K.C.

TITLE Direct Submission

JOURNAL Submitted (07-MAR-2000) Botany, National Taiwan University, 1

FEATURES

source

1..777

/organism="Escherichia coli"

/strain="CH-C2"

/serotype="O-136"

/specific_host="pig"

/db_xref="taxon:562"

/country="Taiwan"

/note="Isolated from hemolytic cells"

1..777

/note="LtpA"

/codon_start=1

/product="heat-labile enterotoxin subunit A precursor"

/protein_id="AA155671.1"

/translation="MKNTFFIFILLASPLVANGDRLYRADSRPDEIKRSGGLMPRG

HNEYFDRTGOMNINLYDHAQRTGTFVRYDDVYTSLSLRSAHLAGSILSGYSTY

IYIATAPMFMVNDVGLVSPHPEOEVSALGGIPIYSQIYGMVNFVDEIRLHN

REYRDYRYRNINIAPEDEYRAGRPDPHOAREPPIHHHAGGGGNSRTTGTCTN

ETTONLSTLYLEKYSKYKRIFFSDYSEVDYIKIRDEL"

1..54

/product="heat-labile enterotoxin subunit A1"

55..630

637..774

/product="heat-labile enterotoxin subunit A2"

BASE COUNT 255 a 136 c 164 g 222 t

ORIGIN

alignment_scores:

Quality: 1304.00 Length: 240

Ratio: 5.433 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:

US-09-528-682-1 x AF242417 ..

Align seg 1/1 to: AF242417 from: 1 to: 777

seq_name: gb_ba:P307LTA

seq_documentation_block:

```

LOCUS      P307LTA              777 bp      DNA      linear      BCT 24-APR-1996
DEFINITION Plasmid P307 (from E. coli) heat-labile enterotoxin subunit A (lta)
gene, complete cds.
ACCESSION  M35581
VERSION    M35581.1 GI:150458
KEYWORDS   enterotoxin; lta gene.
SOURCE     Plasmid P307 (clone: pAT151.) DNA.
ORGANISM   Plasmids.

```

Plasmids.

REFERENCE 1 (bases 1 to 777)
 AUTHORS Dykes, C.M., Halliday, I.J., Hobden, A.N., Read, M.J. and Harford, S.
 TITLE A comparison of the nucleotide sequence of the A subunit of heat-labile enterotoxin and cholera toxin
 JOURNAL FEMS Microbiol. Lett. 26, 171-174 (1985)
 FEATURES
 source 1. 777
 /organism="Plasmid P307"
 /plasmid="plasmid P307"
 /specific_host="Escherichia coli"
 /db_xref="taxon:2472"
 /clone="PAT153."

gene
 1. 777
 /gene="lta"

CDS
 1. 777
 /gene="lta"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin subunit A"
 /protein_id="AAA98202.1"
 /db_xref="GI:150459"

BASE COUNT 255 a 136 c 164 g 222 t
 ORIGIN

alignment_scores:
 Quality: 1304.00 Length: 240
 Ratio: 5.433 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x P307LTA ..

Align seg 1/1 to: P307LTA from: 1 to: 777

```

1  AaNgIyAsPaRgLeuTyArGaAlaAsPSeArGProProAsPgluIleLy 17
  |||||||
55  AATGGCGCACAGATTATACCGTCTGACTGACCCCAAGATGAATAAA 104
  |||||||
17  sArGSeRgLyLeuMeRProArGlyHIsaNgLuTyPhAsPrg 34
  |||||||
105  ACGTTCGGAGGTCTTATGCCAGAGGCGCAATAGACTTCTCGATGAG 154
  |||||||
34  lYThrGlMeTaEnIleAsnLeuTyArPhIsAlaArGlyThrGlThr 50
  |||||||
155  GAACCTCAATGATATTAATCTTATGATCAGCGAGAGGACACAAACC 204
  |||||||
51  GlYPhEValArGtyrAsPaRgLyTyRValSerThrSerLeuSerLeuAr 67
  |||||||
205  GCGTTTGTCAGATATGATGACGATATGTTCCACTCTCTAGTTGAG 254
  |||||||
67  gSeRAlaHIsleuAlaGlyInSerIleLeuSerGlyTyRSerThrTyT 84
  |||||||
255  AAGGCGCACTTAGCAGACAGTCTATATATACAGATATTCACCTACT 304
  |||||||
84  yIlleTyValIleAlaThrAlaProAsnMeRPhEaEnValAsnAsPVal 100
  |||||||
305  ATATATATGTATAGCGACAGCAACAAATATGTTTAATGATGATGA 354
  |||||||
101  LeuGlYValTySerProHIsProTyRGlInGluIleValSerAlaLeuG 117
  |||||||
355  TTAGGCGCTATACAGCCCTCACCATATGACAGAGAGGTTCTCGGTTAG 404
  |||||||
117  yGlyIleRProTyRSeRgInIleTyRgLyTyRPyArGaValAsnPhEgLy 134
  |||||||
405  TGGAAATACCAATATCTCAGATATATGATGATGATCGGTTAAATTTG 454
  |||||||
134  AlIleAsRgIuaRgLeuHIsArGAsnArGsluTyArGaRPaRgTyTyT 150
  |||||||

```

seq_name: gb_pat:A04913

seq_documentation_block:

LOCUS A04913 777 bp DNA linear PAT 13-JUL-1993
 DEFINITION E.coli LTA gene for labile toxin A.
 ACCESSION A04913
 VERSION A04913.1 GI:412215
 KEYWORDS labile toxin A.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1 to 777)
 AUTHORS Hayes, M.V., Harford, S. and Ross, G.W.
 TITLE Microbiological process
 JOURNAL Patent: EP 0145486-A 5 19-JUN-1985;
 GLAXO GROUP LIMITED

FEATURES
 location/Qualifiers

source

CDS

1. 777
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1. 777
 /codon_start=1
 /transl_table=11
 /product="labile toxin A"
 /protein_id="CAA00402.1"
 /db_xref="GI:412216"
 /db_xref="SWISS-PROT:P06717"
 /translation="MKNTFFIFFILLASPLYANGDRLYRADSRPPDEIKRSGLMPRG
 HNEFYDRGTQMNINLYDHARGTGTGFVRYDDGYSTLSLSRAHLAAGSILSGYSTYY
 IYVATAPNMFVNVDVLGVYSPHPYEOEVSALGIPYQIYGMVYVNFVGLDERLHRN
 REYRDYRYRNLIAPAEDGYRLAGPPPHOAMREBPWIHHAPOGCGNSSRITGDTCN
 EETONLSRTYLYREYOSKYKROIIFS DYOSEVDIYNIRDEL"

BASE COUNT 255 a 136 c 164 g 222 t
 ORIGIN

alignment_scores:
 Quality: 1304.00 Length: 240
 Ratio: 5.433 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x A04913 ..

Align seg 1/1 to: A04913 from: 1 to: 777

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluIleu 17
 55 AATGGCGACAGATTATACCGTCTGACTAGACCCCGACATATAATAA 104
 17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 105 ACCTTCGGAGGCTTATGCCAGAGGACATATAGTACTTCGATAGAG 154
 34 LyrhrcInMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 155 GAACCTCAATGAAATATTAACTTTATGATCAGCGAGAGAACACAACC 204
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 205 GCCTTTGTCAGATATGATGACGATATGTTCCACTTCCTTGTAG 254
 67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
 255 AAGTGTCACTTAGCAGACAGCTATATTATCAGATATTCCACTTACT 304
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 305 ATATATATGTTATAGCGACAGACCAAAATATGTTAATGTTAATGAT 354
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
 355 TTAGGGGTATACAGCCCTCCATCATGAAACAGAGGTTCTCGTTAG 404
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
 405 TGGAAATACCATTTCTCAGATATATGATGATGTTATCGTTAATTTG 454
 134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
 455 TGGTTGATGACGATTACATCGTACAGGGAATATAGACCGGATATAC 504
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 505 AGAATCTGATATAGTCCGCGCAGAGGATGTTACAGATTACAGGTTT 554
 167 eProProAspHisGlnAlaTyrArgGluGluProTyrPheHisAlaP 184
 555 CCCACCGGATACCAAGCTTGAGAGAGAACCTGATTCATCATGACAC 604
 184 rGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAs 200
 605 CACAGGTTGTGAAATTCATCAAGAACATACAGGTGATCTGTAAT 654
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
 655 GAGGAGACCCAGAACTGAGCACAATATATCTCAGGGAATATCAATCA 704
 217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 705 AGTTAAGAGCAGATATTTTCAGACATATCAGTACAGGTTGACATATA 754
 234 snArgIleArgAspGluLeu 240
 755 ACAGATTCGGGATGAATTA 774

seq_name: qb_ba: ENSLTPA

seq_documentation_block:

LOCUS ENSLTPA 866 bp DNA linear BCT 26-APR-1993
 DEFINITION Plasmid ENT P307 (enterotoxigenic E.coli) LTP subunit A gene.
 ACCESSION M15362
 VERSION M15362.1 GI:148337
 KEYWORDS LTP gene; heat-labile toxin.
 SOURCE Plasmid ENT (enterotoxigenic E.coli porcine isolate) DNA, clone
 P370.
 ORGANISM Plasmid ENT
 plasmids.

REFERENCE 1 (bases 1 to 866)
 AUTHORS Yamamoto,T., Gojobori,T. and Yokota,T.
 TITLE Evolutionary origin of pathogenic determinants in enterotoxigenic
 JOURNAL Escherichia coli and Vibrio cholerae O1
 MEDLINE J. Bacteriol. 169, 1352-1357 (1987)
 COMMENT Clean copy of sequence [1] kindly provided by T.Yamamoto
 (22-APR-1987).
 FEATURES
 source location/Qualifiers
 1. 866
 /organism="Plasmid ENT"
 /plasmid="Plasmid ENT"
 /db_xref="taxon:2463"
 sig_peptide
 90.143
 /note="Heat-labile toxin subunit A signal peptide"
 CDS
 90..866
 /note="Heat-labile toxin subunit A precursor"
 /codon_start=1
 /transl_table=11
 /protein_id="AAA24793.1"
 /db_xref="GI:148338"
 /translation="MKNTFFILFLASPLYANGDRLRADSRPDEIKRSGLMRPG
 HNEFDRGTQKINIMLKHARCTQGPVRYDGVYSLSLSASHLAAGSISGYSY
 IYVATAPNMFNVDVGVSPHPEEVSALGIPYSQITGRVNFYIDELAHN
 REYRDYRNLNLAIPADGVRLAGFPEDHQAAREPWTHAPDQCGNSSRTTSDTCN
 EETONLSYIYREVQSKVKRQIFSDYCEVDIYRIRDEL"
 mat_peptide
 144..863
 /product="Heat-labile toxin subunit A"
 BASE COUNT 277 a 149 c 181 g 259 t
 ORIGIN Unreported.
 alignment_scores:
 Quality: 1304.00 Length: 240
 Ratio: 5.433 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-528-682-1 x ENSLTPA ..
 Align seg 1/1 to: ENSLTPA from: 1 to: 866
 1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluIleu 17
 144 AATGGCGACAGATTATACCGTCTGACTAGACCCCGACATATAATAA 193
 17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 194 ACCTTCGGAGGCTTATGCCAGAGGACATATAGTACTTCGATAGAG 243
 34 LyrhrcInMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 244 GAACCTCAATGAAATATTAACTTTATGATCAGCGAGAGAACACAACC 293
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 294 GCCTTTGTCAGATATGATGACGATATGTTCCACTTCCTTGTAG 343
 67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
 344 AAGTGTCACTTAGCAGACAGCTATATTATCAGATATTCCACTTACT 393
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 394 ATATATATGTTATAGCGACAGCAAAATATGTTAATGTTATGATGAT 443
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
 444 TTAGGGGTATACAGCCCTCCATCATGAAACAGAGGTTCTCGTTAG 493
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
 494 TGGAAATACCATTTCTCAGATATATGATGATGTTATCGTTAATTTG 543

134 a1lleaspgluargleuHisargasnargluuTyraAspargTyrr 150
|||||
544 TGATGTGTAACGATTTCATCGTACAGGAAATATAGACCGGTATTAC 593
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrrArgLeuAlaGlyPh 167
|||||
594 AGAAATCTGAATATAGCTCCGGCAGAGATGTTACAGATTAGAGGTTT 643
|||||
167 ePrProAspHisGlnAlaIleTrrParGluGluProTrrIleHisAlaP 184
|||||
644 CCACCCGAGTACCCACCACTTGAGAGAAACCTGGATTATCATCATCAC 693
|||||
694 CACAGGTTTGGAAATTCATCAAGAACATACACAGGTGATTCTTGAAT 743
|||||
201 GluGluThrGlnAsnLeuSerThrIleTyrrLeuArgGluuTyrrGlnSer 217
|||||
744 GAGAGACCCAGAAATCTGACACAAATATCTCAGGAAATATCAATCAA 793
|||||
217 sValIysArgGlnIlePheSerAspTyrrGlnSerGluValAspIleTyrr 234
|||||
794 AGTTAAGACGACAGTATTTTCAGACTATCAGAGTTGACATATATA 843
|||||
234 snArgIleArgAspGluLeu 240
|||||
844 ACAGAAATTCGGGATGAATTA 863
|||||
seq_name: gb_ba:ENRLTPA

seq_documentation_block:
LOCUS ENRLTPA 898 bp DNA linear BCT 26-APR-1993
DEFINITION Plasmid ENT-R pCC86 (enterotoxigenic E.coli) LTP subunit A gene.
ACCESSION M15361
VERSION M15361.1 GI:148333
KEYWORDS LTP gene; heat-labile toxin.
SOURCE Plasmid ENT-R (enterotoxigenic E.coli porcine isolate) DNA, clone pCC86.
ORGANISM Plasmid ENT-R
plasmids.
REFERENCE 1 (bases 1 to 898)
Yamamoto,T., Gojobori,T. and Yokota,T.
Evolutionary origin of pathogenic determinants in enterotoxigenic
TITLE Escherichia coli and vibrio cholerae O1
JOURNAL J. Bacteriol. 169, 1352-1357 (1987)
MEDLINE 87137303
COMMENT Clean copy of sequence [1] kindly provided by T.Yamamoto
(22-Apr-1987).

FEATURES
source location/Qualifiers
1..898
/organism="Plasmid ENT-R"
/plasmid="Plasmid ENT-R"
/db_xref="taxon:2462"
122..175
/note="Heat labile toxin subunit A signal peptide"
122..898
/note="Heat labile toxin subunit A precursor"
/codon_start=1
/transl_table=11
/protein_id="AAA24791.1"
/db_xref="GI:148334"
/translation="MKNTFFILLASPIYANGRLRADSRPPDEIKRSGLMPRG
HNEYFDRTOMKINILYDHAQRTQTFVRIDGVSTSLKSAHLAAGSILSGYSTYY
IVIVATAPNMFENVNDVGVISPHPEQEVSAAGGIPYQIYGVYVNFVGIIDERLHRN
REYDRYVRLNLIAPAEADGVRLAGFPDPHOAWREBPWIIHAFQGGNSRFTYGDTCN
EETONLSTIYLYREYSKYVKRQIFSDQSEVDIYNRIDEL"
176..895
/product="Heat labile toxin subunit A"

mat_peptide
BASE COUNT 289 a 155 c 106 g 268 t
ORIGIN Unreported.

alignment_scores:
Quality: 1304.00 Length: 240
Ratio: 5.433 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x ENRLTPA ..
Align seq 1/1 to: ENRLTPA from: 1 to: 898

1 AsnGlyAspArgLeuTyrrArgAlaAspSerArgProProAspGluIleTy 17
|||||
176 AATGGCAGACAGATTATACCGGTGCTGACTGTAGACCCAGATGAATAAA 225
|||||
17 sArgSerGlyGlyLeuMetProAlaGlyHisAsnGluTyrrPheAspArg 34
|||||
226 ACGTTCGGAGGCTCTTATGCCCAAGGGCATTAAAGACTACTTCATAGAG 275
|||||
34 IYThrGlnMetAsnIleAsnLeuTyrrAspHisAlaIleArgGlyThrGlnThr 50
|||||
276 GAACCTCAAATGAATATTATATCTTTATGATCACCGGAGAGAACCAACC 325
|||||
51 GlyPheValArgTyrrAspAspGlyTyrrValSerThrSerLeuSerLeuAr 67
|||||
326 GCGTTTGCAGATATGATGAGCGGATATGTTCCACTTCTGTACTTGGAG 375
|||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrrSerThrTyrr 84
|||||
376 AAGGCTCACCTTACAGACAGCTTATATTATCAGGATATTCACACTTACT 425
|||||
84 YrIleTyrrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
426 ATATATATGTTATATAGCCAGACACCAATATGTTAAATGTTAATGATGA 475
|||||
101 LeuGlyValTyrrSerProHisProTyrrGluGlnGluValSerAlaLeuG 117
|||||
476 TTAGGCGTATACAGCCCTCACCATATGAAACAGAGGTTTCTGGCTTAGG 525
|||||
117 YGlyIleProTyrrSerGlnIleTyrrGlyTyrrTyrrArgValAsnPheGly 134
|||||
526 TGGAAATACCAATATCTCAGATATGATGTATACGTGTTAATTTTGGTG 575
|||||
134 a1lleaspgluargleuHisargasnargluuTyraAspargTyrr 150
|||||
576 TGATGTGTAACGATTTCATCGTACAGGGAATATAGACCGGTATTAC 625
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrrArgLeuAlaGlyPh 167
|||||
626 AGAAATCTGAATATAGCTCCGGCAGAGATGTTACAGATTAGCAGGTTT 675
|||||
167 ePrProAspHisGlnAlaIleTrrParGluGluProTrrIleHisAlaP 184
|||||
676 CCACCCGAGTACCCACCACTTGAGAGAAACCTGGATTATCATCATCAC 725
|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
726 CACAGGTTTGGAAATTCATCAAGAACATACACAGGTGATTCTTGAAT 775
|||||
201 GluGluThrGlnAsnLeuSerThrIleTyrrLeuArgGluuTyrrGlnSer 217
|||||
776 GAGAGACCCAGAAATCTGACACAAATATCTCAGGAAATATCAATCAA 825
|||||
217 sValIysArgGlnIlePheSerAspTyrrGlnSerGluValAspIleTyrr 234
|||||
826 AGTTAAGACGACAGTATTTTCAGACTATCAGTACAGAGTTGACATATATA 875
|||||
234 snArgIleArgAspGluLeu 240
|||||
876 ACAGAAATTCGGGATGAATTA 895
|||||
seq_name: gb_pat:A04915
seq_documentation_block:

LOCUS A04915 777 bp DNA linear PAT 14-JUL-1993
 DEFINITION Artificial sequence LTA gene for labile toxin A.
 ACCESSION A04915
 VERSION A04915.1 GI:412519
 KEYWORDS labile toxin A.
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 777)
 AUTHORS Hayes, M.V., Hartford, S. and Ross, G.W.
 TITLE Microbiological process
 JOURNAL Patent: EP 0145486-A 7 19-JUN-1985;
 GLAXO GROUP LIMITED
 FEATURES
 source Location/Qualifiers
 1..777
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..777
 /codon_start=1
 /transl_table=11
 /product="labile toxin A"
 /protein_id="CAA00403.1"
 /db_xref="GI:412520"
 /translation="MNITFTFILLASPLVANGRLYRADSRPPDEIKRSGCLMPRG
 HNEYPDRGTOMNINIDYDARGQTFEVRVDGYVXTSLSRSAHLAAGOSILSGYSTY
 IYVATAPNMFVNDVLYGVSHPYEQEYSAAGLIPYSQIYGMVHNVGVIDERLHRN
 REYRDYRVNLIITAPEDGYRLAGFPDHOAWREBPWHIHAQGGNSRITGTPTCN
 EETONLSRTIYLAEVQSKVKROIFESDYOSVDIYNRIDEL"
 BASE COUNT 255 a 135 c 164 g 222 t 1 others
 ORIGIN
 alignment_scores:
 Quality: 1300.00 Length: 240
 Ratio: 5.417 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.583
 alignment_block:
 US-09-528-682-1 x A04915 ..
 Align seg 1/1 to: A04915 from: 1 to: 777
 1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
 |||||
 55 AATGGCGACAGATATACCGTGCATGACCCCAATGAATGAA 104
 |||||
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 |||||
 105 ACGTTCGCGAGGCTCTATGCCAGAGGCGCAATGAGTACTTCGATGAG 154
 |||||
 34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 |||||
 155 GAATCGAAATGAATATATATCTTATGATCATCGGAGAGACCAAC 204
 |||||
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 |||||
 205 GCGTTTGTGAGATGATGACGAGATATGTTTTCACCTCTCTAGTTGAG 254
 |||||
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
 |||||
 255 AAGGCTCCTACTTACGAGACAGCTATATATTCAGGATATTCCTACT 304
 |||||
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||
 305 ATATATATGTTATGCGACACCAATATGTTTAAATGTAATGATGTA 354
 |||||
 355 TTAGCGGTATAGACCCCTCACCATATGACAGAGAGGTTTCGCGTAGG 404
 |||||
 117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
 |||||
 405 TGGAAATACCATATTCGATATATGATGATGATGATGATGATGATGATG 454
 |||||

seq_name: gb_pat:E00613
 seq_documentation_block:
 LOCUS E00613 777 bp DNA linear PAT 29-SEP-1997
 DEFINITION DNA encoding A component of heat labile toxin, LTA.
 ACCESSION E00613
 VERSION E00613.1 GI:2168892
 KEYWORDS JP 1986005097-A/1.
 SOURCE JP 1986005097-A/1.
 ORGANISM Escherichia coli.
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 777)
 REFERENCES
 AUTHORS Maikuru, B. H., Suteilbun, H. and Goodon, D. R.
 MICROBIAL METHOD
 JOURNAL Patent: JP 1986005097-A 1 10-JAN-1986;
 GLAXO GROUP LTD
 COMMENT
 OS Escherichia coli
 PN JP 1986005097-A/1
 PD 10-JAN-1986
 PF 12-DEC-1984 JP 1984262645
 PR 12-DEC-1983 GB 83 8333131
 PI MAIKURU BIKUTAA HEIZU, SUTEILIBUN HAAFUODOO,
 PI GOODON UIRAMU ROSU
 PC CO7K15/04, A61K39/108, C12N1/20, C12N15/00//C12P21/00, (C12N1/20,
 PC C12R1:19),
 PC (C12N15/00, C12R1:19);
 CC strandedness: Single;
 CC topology: linear;
 CC hypothetical: No;
 CC anti-sense: No;
 FH key
 CDS 1..777
 FT 1..777
 FT mat_peptide 1..774
 FT toxin, LTA
 FT /product="A component of heat labile toxin, LTA"
 FT mutation
 FT /note="SDMT"
 FT Location/Qualifiers
 1..777
 /organism="Escherichia coli"

BASE COUNT 255 a 137 c 163 g 222 t
 ORIGIN

alignment_scores:
 Quality: 1298.00 Length: 240
 Ratio: 5.408 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.167

alignment_block:
 US-09-528-682-1 x E00613 ..

Align seg 1/1 to: E00613 from: 1 to: 777

```

1 ASngLYAspArgLeuTyrArgAlaAspSerArgProProAspGluIleu 17
|||||
55 AATGGCGACAGATTATATACCTGCTGACTGACACCCCGAGATGAATAAA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
105 ACGTTCGGAGGCTTATGCCCCAGAGGCAATATAGTACTTCGATPAGAG 154
34 lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
155 GAACCAATGAATATATATCTTTATGATCACCAGAGAACACAAACC 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||
205 GGCTTTGTCAGATATGATGACGAGATATGTTCCACTTCTTAGTTGAG 254
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||
255 AAGTGTCACTTATGACGAGAGCTTATATATATCAGATATATCCACTTAC 304
84 yrlIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
305 ATATATATGTTATAGCGACAGACCAATATATGTTATATGTTATGATGTA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||
355 TTAGGGGTATACAGCCCTCACCATATGAAACAGAGTTTCTCGTTAGAG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGlyV 134
|||||
405 TGGAAATACCATATTTCTCAGATATATGATGATGATCGTTAATTTTGGTG 454
455 TGATTGATGACAGATTACATCGTAACAGGAATATAGAGACCGGATATTAC 504
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
|||||
505 AGAAATCTCAATATATAGCTCCGCGACAGAGATGTTACAGATTACAGAGTT 554
167 eProProAspHisGlnAlaThrArgGluIleProTyrPheHisAlaP 184
|||||
555 CCCACCGGATCCGAAAGCTTGAGAGAGAACCTTGATTCATCATGACAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
605 CACAAAGGTTGTGAATTCATCAAGAAACATCACAGGTGATCTACTGTAA 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
|||||
655 GAGGAGACCCAGATCTGAGCACAAATATATCTCAGGGAATATCAATCAAA 704
217 sValTyrArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||
705 AGTTAAGAGGACAGATATTTTTCAGACTATACAGTACAGAGTTGACATATA 754
234 snArgIleArgAspGluLeu 240
|||||

```

755 ACAGAAATTCGCGATGAATTA 774

seq_name: gb_ba:EWDLTA

seq_documentation_block:

LOCUS EWDLTA 723 bp DNA linear BCT 23-APR-1996
 DEFINITION Plasmid EMD 299 (from E.coli) heat-labile enterotoxin A
 (LT-A) gene, 3' end.
 ACCESSION M57244 J05702
 VERSION M57244.1 GI:531190
 KEYWORDS enterotoxin; heat-labile enterotoxin.
 SOURCE Plasmid EMD 299 DNA.
 ORGANISM plasmids.

REFERENCE 1 (bases 1 to 723)
 AUTHORS Tsujii,T., Inoue,T., Miyama,A., Okamoto,K., Honda,T. and Miwatani,T.
 TITLE A single amino acid substitution in the A subunit of Escherichia
 coli enterotoxin results in a loss of its toxic activity
 JOURNAL J. Biol. Chem. 265 (36), 22520-22525 (1990)
 COMMENT 91093102
 FEATURES
 source
 1..723
 location/Qualifiers
 /organism="Plasmid EMD 299"
 /plasmid="Plasmid EMD 299"
 /specific_host="Escherichia coli"
 /db_xref="taxon:2464"
 1..723
 /gene="LT-A"
 <1..723
 /gene="LT-A"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin subunit A"
 /protein_id="AAB59161.1"
 /db_xref="GI:148532"
 /translation="NGDRLYRADSNPPDEIKRSGGLMRGHNEXEDRGTOMNINLYDH
 ARGQTGFRYRDDGVSTSLRSALHAGOSILSGSYTYIVLATAPNMNVNDVIG
 VYSPHPIQAKXSALAGIPYSOITYCHRYRNFVIDERLHGNREYRDRYRNINAPAE
 VYRLAGFPDHDQAMREFPWIRHAPQGCNSSRITYTGDCNEETONLSTIYLRYSKY
 KRQIFSDYQSEVDIYINRDEL"
 459

old_sequence

BASE COUNT 238 a 128 c 160 g 197 t
 ORIGIN

alignment_scores:
 Quality: 1293.00 Length: 240
 Ratio: 5.410 Gaps: 0
 Percent Similarity: 99.583 Percent Identity: 99.167

alignment_block:
 US-09-528-682-1 x EMDLTA ..

Align seg 1/1 to: EMDLTA from: 1 to: 723

```

1 ASngLYAspArgLeuTyrArgAlaAspSerArgProProAspGluIleu 17
|||||
1 AATGGCGACAGATTATATACCTGCTGACTGACACCCCGAGATGAATAAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
51 ACGTTCGGAGGCTTATGCCCCAGAGGCAATATAGTACTTCGATPAGAG 100
34 lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
101 GAACCAATGAATATATATCTTTATGATCACCAGAGAGAACACAAACC 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||

```



```

|||||
151 GGGTTTGTCAGATATGATGACGGATAGTGTCCACTCTCTGATTGAG 200
67 gserAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||
201 AAGTGCACCTTAGCAGACAGCTATATATATCAGGATATCCACTACT 250
84 ylleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
251 ATATATATGTATATAGCCACACGCAAAATATGTTAATGTTAATGATGA 300
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||
301 TTAGGCGGTATACAGCCCTCACCATATGACAGAGAGGTTCTGCGTTAG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhgGly 134
|||||
351 TGGAAATACATATTCCTCAGATATATGATGATACGTGAATTTTGGTG 400
134 aAlleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
401 TCGATTGATGAACGATTTCATGTAACAGGAAATATAGAGACCGGATATAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
451 AGAAATCTGAATATAGCTCCGCGCAGAGATGTTACAGATTACAGAGTTT 500
167 eProProAsnHisGlnAlaThrPargGluGluProTyrIleHisHisAlaP 184
501 CCCACCGGATCCACCAAGCTTGAGAGAAACCTGGATTCAATCATGTCAC 550
184 rGlnGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
551 CACAGAGTTGTGAAATTCATCAAGAAATCATCACAGGTGATCTTGAAAT 600
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIy 217
601 GAGGAGACCCAGAAATCGACACATATATCTCAGAGGAATATCAATCAA 650
217 sValIlyArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
651 AGTTAAAGGACAGATATTTTCAACCTATTCAGTCAGAGGTTGACATATA 700
234 snArgIleArgAspGluLeu 240
701 ACAGAAATTCGGAGTGAATTA 720
seq_name: gb_pat:E03421
seq_documentation_block:
LOCUS E03421 1148 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding Lth of entero toxigenic Escherichia coli.
ACCESSION E03421.1 GI:2171637
VERSION E03421.1 JP 1992079898-A/1.
KEYWORDS JP 1992079898-A/1.
SOURCE Escherichia coli.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1 (bases 1 to 1148)
Danbara,H. and Abe,A.
DNA AND RNA PROBE TO SIMULTANEOUSLY DETECT VIBRIO CHOLERAE AND
ENTEROTOXIGENIC ESCHERICHIA COLI AND DETECTION OF VIBRIO CHOLERA
AND ENTEROEXIGENIC ESCHERICHIA COLI USING THE SAME DNA AND RNA
PROBE
Patent: JP 1992079898-A 1 13-MAR-1992;
JOURNAL
KITASATO INST:THE
OS Escherichia coli
PN JP 1992079898-A/1
PD 13-MAR-1992
PF 23-JUL-1990 JP 1990194208
PI DANBARA HIROFUMI, ABE AKIO
PC C1201/68,C07H21/02,C07H21/04,C12N15/11,C12N15/31,C1201/04, PC

```

```

C1201/10,
PC (C12N15/31,C12R1:19),(C1201/04,C12R1:63,C12R1:19): CC
strandedness: Double:
CC topology: linear:
CC hypothetical: No:
CC anti-sense: No:
CC *source: clone-pkAD008;
FH key Location/Qualifiers
FH misc-feature 1..1148
FT /note="DNA fragment encoding Lth of entero FT
FT sig_peptide 1..54
FT sig_peptide 55..774
FT mat_peptide 1..777
FT CDS /product="subunit A of Lth"
FT sig_peptide 774..837
FT /product="precursor subunit A of Lth" FT
FT mat_peptide 838..1145
FT /product="subunit B of Lth"
FT CDS 774..1148
FT /product="precursor of subunit B of Lth"
FEATURES
source Location/Qualifiers
1..1148
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 399 a 205 c 229 g 315 t
ORIGIN
alignment_scores:
Quality: 1292.00 Length: 240
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.750
alignment_block:
US-09-528-682-1 x E03421 ..
Align seg 1/1 to: E03421 from: 1 to: 1148
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
|||||
55 AATGGCGACAAATTTATACCGTCTGACTCTACACCCCGAGTCAATATAA 104
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
105 ACGTCCGGAGGCTTATATGCCAGAGGCAATATGAGTCTTCGATAGAG 154
34 lYthrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
155 GAATCTCAATAGATATTAATCTTTATGATCAGCCGAGAGAACACAAACC 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
205 GCCTTTGTCAGATATGATGACGGATATGTTCCACTTCCTCTGATTGAG 254
67 gserAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||
255 AAGTGCCTCACTTAGCAGACAGCTATATATATCAGATATTCACACTTACT 304
84 ylleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
305 ATATATATGTATATAGCAGACAGCAAAATATGTTAATGTTAATGATGA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||
355 TTAGGCGGTATACAGCCCTCACCATATGAAACAGAGAGGTTCTGCGTTAG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhgGly 134
|||||
405 TGGAAATACATATTCCTCAGATATATGATGATACGTGAATTTTGGTG 454

```


134 allleaspgluargleuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
455 TGATGTGTAAGCATTTACATCGTATACAGGAAATATAGACCGGATATAC 504
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||
505 AGAAATCTGATATAGCTCCGGCAGAGATGTTACAGATATACAGAGTTT 554
|||||
167 eProProAspHisGlnAlaTPrArgGluGluProTPrIleHisHisAlaP 184
|||||
555 CCCACCGGATCCACCACTTGAGAGAGAACCCTGGATTCATATATGCAC 604
|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
605 CACAGAGTTGTGGAATTCATCAGAAGACATTTACAGGTGATCTGTAT 654
|||||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLy 217
|||||
655 GAGGAGACCCGAAATCTGACACACAAATATCTCAGAAATATCAATCAAA 704
|||||
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||
705 AGTTAAGAGCAGATATTTTCAGACTATGCTCAGAGGTTGACATATATA 754
|||||
234 snArgIleArgAspGluLeu 240
|||||
755 ACAGAAATCGGAATGATTA 774

seq_name: gb_ba:AB011677

seq_documentation_block:

LOCUS AB011677 1434 bp DNA linear BCT 23-APR-1998
DEFINITION Escherichia coli genes for heat-labile enterotoxin A subunit and B
subunit, complete cds.

ACCESSION AB011677.1 GI:3062900
VERSION heat-labile enterotoxin B subunit; Lth b subunit; heat-labile
enterotoxin A subunit; Lth a subunit.

KEYWORDS Escherichia coli (strain:1032 (Enterotoxigenic)) DNA.
Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

SOURCE 1 (bases 1 to 1434)
Komase, K.
Direct Submission

REFERENCE 1 (bases 1 to 1434)
Komase, K.
Submitted (02-MAR-1998) Katsunori Komase, The Kitasato Institute,
Center for Basic Research, Lab. of Virology, Shirokane 5-9-1,
Minato-ku, Tokyo 108-8642, Japan (E-mail:komase-kek@kitasato.or.jp,
Tel:+81-3-3444-6161, Fax:+81-3-3444-6637)

JOURNAL 2 (sites)
Tanura, S., Asanuma, H., Tomita, T., Komase, K., Kawahara, K.,
Danbara, H., Hattori, N., Watanabe, K., Suzuki, Y., Nagamine, T.,
Alizawa, C., Oya, A. and Kurata, T.
Escherichia coli heat-labile enterotoxin B subunits supplemented
with a trace amount of the holotoxin as an adjuvant for nasal
influenza vaccine

REFERENCE 1 (bases 1 to 1434)
Komase, K.
Submitted (02-MAR-1998) Katsunori Komase, The Kitasato Institute,
Center for Basic Research, Lab. of Virology, Shirokane 5-9-1,
Minato-ku, Tokyo 108-8642, Japan (E-mail:komase-kek@kitasato.or.jp,
Tel:+81-3-3444-6161, Fax:+81-3-3444-6637)

JOURNAL 2 (sites)
Tanura, S., Asanuma, H., Tomita, T., Komase, K., Kawahara, K.,
Danbara, H., Hattori, N., Watanabe, K., Suzuki, Y., Nagamine, T.,
Alizawa, C., Oya, A. and Kurata, T.
Escherichia coli heat-labile enterotoxin B subunits supplemented
with a trace amount of the holotoxin as an adjuvant for nasal
influenza vaccine

REFERENCE 1 (bases 1 to 1434)
Komase, K.
Submitted (02-MAR-1998) Katsunori Komase, The Kitasato Institute,
Center for Basic Research, Lab. of Virology, Shirokane 5-9-1,
Minato-ku, Tokyo 108-8642, Japan (E-mail:komase-kek@kitasato.or.jp,
Tel:+81-3-3444-6161, Fax:+81-3-3444-6637)

JOURNAL 2 (sites)
Tanura, S., Asanuma, H., Tomita, T., Komase, K., Kawahara, K.,
Danbara, H., Hattori, N., Watanabe, K., Suzuki, Y., Nagamine, T.,
Alizawa, C., Oya, A. and Kurata, T.
Escherichia coli heat-labile enterotoxin B subunits supplemented
with a trace amount of the holotoxin as an adjuvant for nasal
influenza vaccine

REFERENCE 1 (bases 1 to 1434)
Komase, K.
Submitted (02-MAR-1998) Katsunori Komase, The Kitasato Institute,
Center for Basic Research, Lab. of Virology, Shirokane 5-9-1,
Minato-ku, Tokyo 108-8642, Japan (E-mail:komase-kek@kitasato.or.jp,
Tel:+81-3-3444-6161, Fax:+81-3-3444-6637)

JOURNAL 2 (sites)
Tanura, S., Asanuma, H., Tomita, T., Komase, K., Kawahara, K.,
Danbara, H., Hattori, N., Watanabe, K., Suzuki, Y., Nagamine, T.,
Alizawa, C., Oya, A. and Kurata, T.
Escherichia coli heat-labile enterotoxin B subunits supplemented
with a trace amount of the holotoxin as an adjuvant for nasal
influenza vaccine

REFERENCE 1 (bases 1 to 1434)
Komase, K.
Submitted (02-MAR-1998) Katsunori Komase, The Kitasato Institute,
Center for Basic Research, Lab. of Virology, Shirokane 5-9-1,
Minato-ku, Tokyo 108-8642, Japan (E-mail:komase-kek@kitasato.or.jp,
Tel:+81-3-3444-6161, Fax:+81-3-3444-6637)

JOURNAL 2 (sites)
Tanura, S., Asanuma, H., Tomita, T., Komase, K., Kawahara, K.,
Danbara, H., Hattori, N., Watanabe, K., Suzuki, Y., Nagamine, T.,
Alizawa, C., Oya, A. and Kurata, T.
Escherichia coli heat-labile enterotoxin B subunits supplemented
with a trace amount of the holotoxin as an adjuvant for nasal
influenza vaccine

REFERENCE 1 (bases 1 to 1434)
Komase, K.
Submitted (02-MAR-1998) Katsunori Komase, The Kitasato Institute,
Center for Basic Research, Lab. of Virology, Shirokane 5-9-1,
Minato-ku, Tokyo 108-8642, Japan (E-mail:komase-kek@kitasato.or.jp,
Tel:+81-3-3444-6161, Fax:+81-3-3444-6637)

gene 864..1238
CDS 864..1238
/gene="Lth b subunit"
/gene="Lth b subunit"
/codon_start=1
/transl_table=11
/product="heat-labile enterotoxin B subunit"
/protein_id="BA25726.1"
/db_xref="GI:3062902"
/translation="MNKVCYVLEFTALLSLICAYGAPOSITELCSFRMTQITINDK
IISYTESAGKREMYITITKSGATFVEYVPGSHIDISOAKAIERMDUTRITITETK
IDKLCVNNKTPNSIAISMEN"

alignment_scores:
Quality: 1290.00 Length: 240
Ratio: 5.397 Gaps: 0
Percent Similarity: 99.583 Percent Identity: 98.750

alignment_block:
US-09-528-682-1 x AB011677 ..

Align seg 1/1 to: AB011677 from: 1 to: 1434

1 AsnGlyAspArgLeuTyrArgAlaSerArgProProAspGluIlely 17
|||||
145 AATGGCGCAATATTAATACCGTGTGACTGTAGACCCGACATGAATATA 194
|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
195 ACGTTCCGAGGCTGTATGCCAGAGGCGCATATAGACTTCGATGAG 244
|||||
34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
245 GACCTCAATGATATTAATCTTATGATCAGCGAGGAGAACCAACC 294
|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
|||||
295 GCGTTTGTCAATATGATGACGGAATGTTTCCACTTCTTACGTTGAG 344
|||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||
345 AAGTGCACACTTACAGAGACGCTATATTTATCAGGAATTCACACTTACT 394
|||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
395 ATATATATGTTATAGCAGACACCAATATATGTTAATGTTAAAGATGA 444
|||||
101 LeuGlyValTyrSerProHisArgTyrGlnGlnGluValSerAlaLeuG 117
|||||
445 TTAGCGGTATACACCCCTCACCCTATATGACAGAGAGGTTTGGCTTAG 494
|||||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
|||||
495 TGAATATGCAATATCTCAGATATATGATGCTATCGTTAATTTTGGTG 544
|||||
134 allleaspgluargleuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
545 TAAATGATGAACGATTTACATCGTAAACAGGGAATATAGACCGGATATTAC 594
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||
595 AGAAATCTGATATAGCTCCGGCAGAGATGTTTACAGATATAGCAGGTTT 644
|||||
167 eProProAspHisGlnAlaTPrArgGluGluProTPrIleHisHisAlaP 184
|||||
645 CCCACCGGATCCACCACTTGAGAGAGAACCCTGGATTCATATATGCAC 694
|||||

184 roGInGlycylAsnSerArgThrIleThrGlyAspThrCysAsn 200
 |||||||:|||||||||||||||||:|||||||||
 695 CACAGAGTGTGGAAATTCATCAAGAACATTCACATGATCTGTAT 744
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIy 217
 |||||||:|||||||||||||||||:|||||||||
 745 GAGGAGACCCAGAACTGAGCACAAATATCTCGAGAAATATCAATCAAA 794
 217 sVallyAsrGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
 |||||||:|||||||||||||||||:|||||||||
 795 AGTTAAGAGCAGATATTTTCAGACTATCATGTCAGAGTTGACATATATA 844
 234 snArgIleArgAspGluLeu 240
 |||||||:|||||||||||||||||:|||||||||
 845 ACAGAAATTCGGGATGAAATTA 864

seq_name: gb_ba:ECOTOXA

seq_documentation_block:

LOCUS ECOTOXA 934 bp DNA linear BCT 26-APR-1993
 DEFINITION E.coli toxA gene encoding subunit A of heat-labile enterotoxin.

ACCESSION K01995

VERSION K01995.1 GI:148027

KEYWORDS enterotoxin; heat-labile enterotoxin; toxA gene.

SOURCE E.coli H10407 (serotype 078:H11) plasmid DNA, clone pJy27.

ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE 1 (bases 1 to 934)

AUTHORS Yamamoto,T., Tamura,T. and Yokota,T.

TITLE Primary structure of heat-labile enterotoxin produced by

Escherichia coli pathogenic for humans

J. Biol. Chem. 259, 5037-5044 (1984)

MEDLINE 84185610

COMMENT This heat-labile enterotoxin is pathogenic for humans (LTn).
 Compared in [1] with LTp A and CT A (Ltp-pathogenic for piglets,
 CT-choleera toxin).

FEATURES

source

1..934

/organism="Escherichia coli"

/db_xref="taxon:562"

CDS

158..934
 /note="heat-labile enterotoxin A prepeptide"

/codon_start=1

/transl_table=1

/protein_id="AA24685.1"

/translation="MKNTFFIFLLASPLYANGDKLYRADSRPDEIKRSGLMPRG
 HNEFDKGTOMNINLYDHARGTGFVARYDGYVSTLSLSAHLAGOSILSGSTYY
 IYVATAPNMNVDVGVSPHPYEDQVSLAGIPIYSQIIGWYRVNFGVIDERLHRN
 REYDRYRNINLPAEDGYRLAGEPPDHQAMREPMTHHAPCCGSDSSRITGDTCN
 EETONLSYIRKTKYKSKYKRIQIFSDYQSEVDIVYRIRREL"

sig_peptide 158..211

/note="heat-labile enterotoxin A signal peptide"

mat_peptide 212..787

/product="heat-labile enterotoxin A1"

mat_peptide 794..931

/product="heat-labile enterotoxin A2"

BASE COUNT 298 a 165 c 186 g 285 t

ORIGIN 46 bp upstream of PstI recognition pattern.

alignment_scores:

Quality: 1287.00

Ratio: 5.362

Percent Similarity: 100.000

Length: 240

Gaps: 0

Percent Identity: 98.333

alignment_block:

US-09-528-682-1 x ECOTOXA ..

Align seg 1/1 to: ECOTOXA from: 1 to: 934

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluIleIy 17
 |||||||:|||||||||||||||||:|||||||||
 212 AATGGGACAAATTTATACCGTGCTGACTAGACCCCGAGATGAATAA 261
 17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 |||||||:|||||||||||||||||:|||||||||
 262 ACGTTCCGAGGCTTATGCCAGAGGCGATATAGTACTTCGATAGAG 311
 34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 |||||||:|||||||||||||||||:|||||||||
 312 GAACTCAATTAATTAATTTATGATGATCAGCGAGAGAACACCAAC 361
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 |||||||:|||||||||||||||||:|||||||||
 362 GGCTTGTCAGATATGATGACGATATGTTCCACTTCCTTAGTTGAG 411
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
 |||||||:|||||||||||||||||:|||||||||
 412 AAGTGCTCACTTAGCAGACAGCTATATATATCAGATATTCACACTTACT 461
 84 YrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||||:|||||||||||||||||:|||||||||
 462 ATATATATGTTATAGCAGACAGACCAAAATATGTTATATGATGATA 511
 101 LeuGlyValTyrSerProHisProTyrGlnGluValSerAlaLeuG1 117
 |||||||:|||||||||||||||||:|||||||||
 512 TTAGCGGTATACAGCCCTCACCATATGAACAGAGGTTTCGCGTTAG 561
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
 |||||||:|||||||||||||||||:|||||||||
 562 TGGAAATACCATATATTCAGATATATGATGATGATCTGTTAATTTGGTG 611
 134 allleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 |||||||:|||||||||||||||||:|||||||||
 612 TGATTTATGACAGCATTTACATCGTAAACAGGAAATATGAAACCGGTATAC 661
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 |||||||:|||||||||||||||||:|||||||||
 662 AGAAATCTGAATATATGCTCCGCGAGAGATGGTTACAGATTGACAGTWT 711
 167 eProProAspHisGlnAlaThrArgGluGluProTyrPheHisAlaP 184
 |||||||:|||||||||||||||||:|||||||||
 712 CCCACCGGATCACCAAGCTTGAGAGAGAACCTTGATTCATCATGCAAC 761
 184 roGInGlycylAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 |||||||:|||||||||||||||||:|||||||||
 762 CACAGAGTGTGGAGATTCATCAAGAACATTCACAGTATCTGTAT 811
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIy 217
 |||||||:|||||||||||||||||:|||||||||
 812 GAGGAGACCCAGAACTGAGCACAAATATCTCGAGAAATATCAATCAAA 861
 217 sVallyAsrGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
 |||||||:|||||||||||||||||:|||||||||
 862 AGTTAAGAGCAGATATTTTCAGACTATCATGTCAGAGTTGACATATATA 911
 234 snArgIleArgAspGluLeu 240
 |||||||:|||||||||||||||||:|||||||||
 912 ACAGAAATTCGGGATGAAATTA 931

seq_name: gb_ba:S60731

seq_documentation_block:

LOCUS S60731

DEFINITION heat-labile enterotoxin A subunit, heat-labile enterotoxin B

subunit [Escherichia coli, 21d, Genomic, 1275 nt].

ACCESSION S60731

VERSION S60731.1 GI:408994

KEYWORDS

SOURCE Escherichia coli 21d.

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1 to 1275)
 AUTHORS Inoue,T., Tsuji,T., Koto,M., Imamura,S. and Miyama,A.
 TITLE Amino acid sequence of heat-labile enterotoxin from chicken enterotoxigenic *Escherichia coli* is identical to that of human strain H 10407
 JOURNAL FEMS Microbiol. Lett. 108 (2), 157-161 (1993)
 MEDLINE 93325225
 PUBMED 8486242
 REMARK Genbank staff at the National Library of Medicine created this entry [NCBI glibseq 131491] from the original journal article. This sequence comes from Fig. 2.
 FEATURES
 source Location/Qualifiers
 1..1275
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 68..844
 CDS
 /note="This sequence comes from Fig. 2: conceptual translation presented here differs from translation in publication: Ltc A subunit"
 /codon_start=1
 /product="heat-labile enterotoxin A subunit"
 /protein_id="AAC60440.2"
 /db_xref="GI:18390349"
 /translation="MKNTFFFIILASPLYANGDKLYRADSRPDEIKRSGLMPRG HNEVFDRTQMNLNDHARCTGTGFEVRYDGYSTLSLSAHLAAGOSILSGYSTY IYVATPMFNVNDVGLGVSPHPEYEOESALGIPYSQIYGVNVRNGVIDERLARN REYRDYRNINLIPAEDEGYSRIAGPPPHQARREPMWIIHARQGGDSRRITGGTCN EETQNLSTIYLRKIQSKYKRQIFSDYSQSEVDIYNIRNEL"
 841..1215
 CDS
 /note="This sequence comes from Fig. 2: Ltc B subunit"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin B subunit"
 /protein_id="AAC60441.1"
 /db_xref="GI:408996"
 /translation="MNKVCYVLFPTALLSLCAGAPOSTELCSEYRNTQIYTIINDK ILSYSEAGKRENVITFEKSGATFQVEVPGSQIHDSQKAIERMKDFLRITLYIETK IDLCVNMNKKTPNSIAIYASMEN"
 BASE COUNT 436 a 227 c 248 g 364 t
 ORIGIN
 alignment_scores:
 Quality: 1287.00 Length: 240
 Ratio: 5.362 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.333
 alignment_block:
 US-09-528-682-1 x S60731 ..
 Align seg 1/1 to: S60731 from: 1 to: 1275
 1 ASnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
 |||||:::|||||
 122 AATGGCACAATAATATACCGTGTGACTGATAGACCCACAGATGAATAAA 171
 |||||
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 |||||
 172 ACGTTCCGGAGGCTTATGCCACAGGGCATTAAGACTACTTCATGAG 221
 |||||
 34 LyrHrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 |||||
 222 GAACCTCAATGAAATTAATCTTATGATCAGCGAGAGACCAACACC 271
 |||||
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 |||||
 272 GCGTTGTGAGATGATGATGACGATATGTTCCACTTCCTTACTGAG 321
 |||||
 67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTyrSerThrTyr 84
 |||||
 322 AAGTGTCACTTAGCAGACAGCTGATATATATGATGATATTCACACTACT 371

84 yIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||
 372 ATATATATGTTATGACGACACCAATAATGTTATGTTAATGATGATA 421
 |||||
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
 |||||
 422 TTAGCGCTATACACCCCTCAACCATATGACAGAGAGGCTTCGCGTAC 471
 |||||
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnGly 134
 |||||
 472 TGGATACCATATTCATGATATATGATGATGATGATGATATATTTG 521
 |||||
 134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
 |||||
 522 TGATGTATGAACGATTTACATCGTAAACGGAATATAGACCGGATATAC 571
 |||||
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 |||||
 572 AGAATCTGAATATAGCTCCGCGAGAGAGTGTACAGATTAACAGGTT 621
 |||||
 167 eProProAspHisGlnAlaThrPargGluGluProTyrIleHisAlaP 184
 |||||
 622 CCCACCGGATCACCAACCTTGAGAGAAGAACCTTGATTCATCATCAC 671
 |||||
 184 rGInGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 |||||
 672 CACAGAGTTGGGAGATTCAACGAACATTTACAGGTGATCTTGAT 721
 |||||
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
 |||||
 722 GAGGAGACCGAGATCTGACACACAAATATCTCAGAAATATCAATCAA 771
 |||||
 217 sValIleArgGlnIlePheSerArgTyrGlnSerGluValAspIleTyr 234
 |||||
 772 AGTTAAGAGCAGATATTTTCAGACTATCATGCTCAGAGGTGACATATA 821
 |||||
 234 snArgIleArgAspGluLeu 240
 |||||
 822 ACAGAAATTCGAATGAATTA 841
 |||||
 seq_name: gb_pat:A16419
 seq_documentation_block:
 LOCUS A16419 711 bp DNA linear PAT 06-JUL-1995
 DEFINITION A subunit of a labile toxin LT1.
 ACCESSION A16419
 VERSION A16419.1 GI:641008
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli.
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 711)
 REFERENCE
 1 IMUNOGENIC DETOXIFIED MUTANTS OF CHOLERA TOXIN AND OF THE TOXIN
 LT, THEIR PREPARATION AND THEIR USE FOR THE PREPARATION OF VACCINES
 Patent: WO 93/3202-A 42 08-JUL-1993;
 JOURNAL
 TITLE
 LOCATION/Qualifiers
 1..711
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1..711
 /gene="LT1 subunit A"
 <1..711
 /gene="LT1 subunit A"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA01280.1"
 /db_xref="GI:641009"
 /translation="NGDRLYRADSRPDEIKRSLMPRGNEVFDRTQMNLNDHAR
 KTOGTGFVARDGYSTLSLSAHLAAGOSILSGYSTIYIYVATPMFNVNDVGLGVSP
 HPEYEOESALGIPYSQIYGVNVRNGVIDERLARNREYRDYRNINLIPAEDEGYSRI
 AGPPPHQARREPMWIIHARQGGDSRRITGGTCNEETQNLSTIYLRKIQSKYKRQIFSDY

439 AGAATGTGATATAGCTCCGGCAGAGATGTTACAGATTACGAGTTT 488
 167 eProPAsphigInalArPaRGluGluProTriPleHisAlap 184
 489 CCCACCGGATACCAAGCTTGAGAGAAACCTGGATTTCATCATGCAC 538
 184 roGInGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 539 CACAGAGTTGGAGATTTCATCAGAACATCATCAGAGTACTTGTAAAT 588
 201 GluGluThrGlnAsnLeuSerThrIleTyrIleuArgIuTyrGlnSerly 217
 589 GAGAGACCCAGAAATCTGAGCACAATATCTCAGGAAATTCATCAACAA 638
 217 sVallyArGInIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 639 AGTTAAAGGACATATTTTCAGACTATCAGTCAAGAGTTGACATATATA 688
 234 snArGIIeaTgAspGluLeu 240
 689 ACAGAAATTCGGGATGAATTA 708

seq_name: gb_ba:ECCELTA

seq_documentation_block:

LOCUS ECCELTA 868 bp DNA linear BCT 12-SEP-1993
 DEFINITION E. coli gene elta encoding the A subunit of the heat-labile
 enterotoxin (LT).
 ACCESSION V00275.1 GI:41339
 VERSION V00275.1 GI:41339
 KEYWORDS enterotoxin; signal peptide.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 868)
 REFERENCE 1 (bases 1 to 868)
 AUTHORS Slicer, E.K. and Noble, J.A.
 TITLE Escherichia coli heat-labile enterotoxin. Nucleotide sequence of
 the A subunit gene
 JOURNAL J. Biol. Chem. 257 (10), 5716-5721 (1982)
 MEDLINE 82167425
 FEATURES

source location/Qualifiers
 1..868
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 mRNA 56..>868
 /product="messenger RNA"
 sig_peptide 90..143
 CDS 90..854
 /codon_start=1
 /transl_table=11
 /product="enterotoxin subunit A"
 /protein_id="CAA23532.1"
 /db_xref="GI:41340"
 /translation="MKNTFIPFILLASLVANGDRLVADSPDEIKRSLMPRG
 NEYFDGTMQINILYDHAAGTGTGVRDDGVYSLSIRSHLAGOYLISGYSLEIY
 IVIAMFNVDIVSYSPHRYQEVSAIGIPIYSITGYRNVFVIDERLRNREIR
 DRYRNLIAPAEDEIRLAGFPDHDQAMREEPHHAPOGCDSSRTTGTGDCNEETQ
 NLSTIYLEYDSKVRQLFSDQSEVDIYNRIIRDEL"

BASE COUNT 284 a 142 c 179 g 263 t
 ORIGIN

alignment_scores:
 Quality: 1183.00 Length: 240
 Ratio: 5.189 Gaps: 2
 Percent Similarity: 95.000 Percent Identity: 93.333

alignment_block:
 US-09-528-682-1 x ECCELTA ..

Align seg 1/1 to: ECCELTA from: 1 to: 868

1 AsnGlyAspArgLeuTyrArgAlaSerSerArgProProAspGluIlely 17
 144 AATGCCACAGATTATACCGTGTGACTTGACCCCAAGATGAATATA 193
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
 194 ACGTTTCGGAGTCTTATGCCAGAGT...AATGAGTACTTCGATAGAG 240
 34 lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 241 GAACCTCAATGAATATTAATCTTATGATCAGCGAGAGGAACACAAAC 290
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 291 GCGTTTGCAGATATGATGACGATATGTTTCCACTTCTTACGTTGAG 340
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
 341 AAGTGCCTCACTTACAGACAGTATATATTCAGAGATATCAGTACTA 390
 84 yTleTyrValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
 391 TATATATCGTTATAGCA.....AATATGTTTAATGTTAATGATGTA 431
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
 432 ATTAGCGTATACAGCCCTCCACCATATGACAGAGAGTTTGTGGTTAG 481
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
 482 TGAATACCATATTCGAGATATATGATGTGATCGTGTATTTTGGTG 531
 134 alIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyr 150
 532 TGATTGATGAACGATTCATGCTGAACAGGAAATAGACCGGTATATAC 581
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 582 AGAATATGAAATATAGCTCCGGCAGAGAGATGTTACAGATTAGAGTTT 631
 167 eProPAsphigInalArPaRGluGluProTriPleHisAlap 184
 632 CCCACCGGATACCAAGCTTGAGAGAAACCTGGATTTCATCATGCAC 681
 184 roGInGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 682 CACAAAGTTGTGAGATTTCATCAAGAACATCACAGAGTACTTGTAAAT 731
 201 GluGluThrGlnAsnLeuSerThrIleTyrIleuArgIuTyrGlnSerly 217
 732 GAGAGAGCCAGAAATCTGAGCACAATATCTCAGGAAATTCATCAACAA 781
 217 sVallyArGInIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 782 AGTTAAAGGACAGATATTTTCAGACTATCAGTCAAGAGTTGACATATATA 831
 234 snArGIIeaTgAspGluLeu 240
 832 ACAGAAATTCGGGATGAATTA 851

seq_name: gb_ba:AF452584

seq_documentation_block:

LOCUS AF452584 777 bp DNA linear BCT 14-JAN-2002
 DEFINITION Vibrio cholerae strain 1322-69 CtxA (ctxA) gene, complete cds.
 ACCESSION AF452584
 VERSION AF452584.1 GI:18141148
 KEYWORDS
 SOURCE Vibrio cholerae.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 1 (bases 1 to 777)

AUTHORS Li, M., Chen, Y., Kotetishvili, M., Stine, O.C., Morris, J.G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.
TITLE Genetic Analysis of the Virulence Regions, CTX f prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Sero-group Strains of Vibrio cholerae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Li, M., Chen, Y., Kotetishvili, M., Stine, O.C., Morris, J.G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA
FEATURES
source 1..777
/organism="Vibrio cholerae"
/strain="1322-69"
/db_xref="taxon:666"
/note="Identical sequence found in Vibrio cholerae strains
571-88 (serogroup O105) and 203-93 (serogroup O141)
serogroup: 037"
1..777
/gene="ctxA"
1..777
/gene="ctxA"
/codon_start=1
/transl_table=1
/product="ctxA"
/protein_id="AA65025.1"
/db_xref="GI:18141149"
/translation="MWKIIFFVFFFLSFSFYANDKLYRADSPDEIKOSGGLMPRG
QNEYDRCTQNNINILDHARGTQTFVHHDDGYVSTSLSAHLVGGTILSGHSTY
IYVITAPNMENVDVLAGSPHPDEQEVSAAGIPIYOIGWVHFVHGVDLEQHRN
RGYRBYNSNDIAPADQYGLAGFPPEHPHAPGCGNAPRSMSTCD
EKTOSIGKFLDEQSKVKRQIFSGYOSDIDITHNRKDEL"

BASE COUNT 243 a 129 c 169 g 236 t
ORIGIN

alignment_scores: Quality: 1093.00 Length: 240
Ratio: 4.836 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 82.083

alignment_block:
US-09-528-682-1 x AF452584 ..

Align seg 1/1 to: AF452584 from: 1 to: 777

1 ASnglYAspaRgLeuTYrArGaLaAsSeRaRgProFroAspGluIlely 17
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
55 AATGATGATATAGTTATATCGGCAGATTCAGACCTCTGATGATAATAA 104
17 sArGSeRgLyGlyLeuMePrOaRgLyHIsAsnGluTYrPhesPaRgC 34
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 GCAGTCAGTGGTCTTATGCCAAGACAGCAAGTACTTTCACGCAG 154
34 lyHrHmeTasnlleAsLeuTYrAspHIsAlaRgLyHnGInThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
155 GTCCTCAAAATGATATCACTTATGATCATGCAGAGSAACTCAACG 204
51 gLyPhesAlaRgTYrAspAspGlyTYrValSerThrSerLeuSerleuAr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 GGAATTGTGGCAGATGATGATAGTTCCACCTCATTTAGTTGAG 254
67 gSeRAlaHIsLeuAlaGlyInSerIleLeuSerGlyTYrSerThrTYrT 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
255 AAGTGGCCACTAGTGGTCAAACTATATTGTCTGGCATTTACTTATTT 304
84 yTlIeTYrValIleAlaThrAlaProAsnMeThesAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ATATATATGTTATAGCCATGACCCCAACATGTTTAACTTAATGATGTA 354

101 LeuGlyValTYrSeRProHIsProTYrGluGluGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 TTAGGGGCATACAGTCTCTCATGCAGATGACAGAGAGTTCTGCTTAGG 404
117 yGlyIleProTYrSeRgInIleTYrGlyTYrPYrArGaValAsnPhesGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 TGGGATTTCCCTACTCTCCCAATATATGATGATGATTCAGTTCAATTTGGGG 454
134 allleAspGluArGLeuHIsArGAsnArGlyLurTYrArGaPaRgTYrTYr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
455 TGCTTGATGACAAATTATCATGTAATAGGGGCTACAGAGATATGATTTAC 504
151 ArGAsnAsnIleAlaProAlaGluAspGlyTYrArGLeuAlaGlyPh 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
505 AGTAACTTAGATATTTGCTCCAGCAGCAGATGATTAGATTTGACAGTTT 554
167 eProProAspHIsGlnAlaTrpArGluGluProTYrIleHIsAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
555 CCTCCGAGCATATAGAGCTTGAGGAGAGCCGTGATTCATCATGACAC 604
184 roGInGlyCysGlyAsnSerSerArGThrIleThrGlyAspThrCysAsn 200
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
605 CGCGGGTTGTGGAAATGCTCAAGATCATGATGATTAATCTTGCGAT 654
201 GLuGluThrGlnAsnLeuSerThrIleTYrLeuArGluTYrGlnSerly 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
655 GAAAAAACCAAGTCTAGTGTAATAATCTCTGAGCAATACCAATCTAA 704
217 sValLySArGInIlePheserAspTYrGlnSerGluValAspIleTYrA 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 AGTTAAAGCAATATTTTACAGCTATCATCTGATTTGATATACATTA 754
234 snArGlyleArGAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
755 ATAGAAATTAGATGATTA 774

seq_name: gb_ba:AF463400

seq_documentation_block:
LOCUS AF463400 777 bp DNA linear BCT 31-JAN-2002
DEFINITION Vibrio cholerae strain 571-88 CtxA (ctxA) gene, complete cds.
ACCESSION AF463400
VERSION AF463400.1 GI:18448888
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Vibrio cholerae.
REFERENCE
1 (bases 1 to 777)
AUTHORS Li, M., Chen, Y., Kotetishvili, M., Stine, O.C., Morris, J.G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.
TITLE Genetic Analysis of the Virulence Regions, CTX f prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Sero-group Strains of Vibrio cholerae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Li, M., Chen, Y., Kotetishvili, M., Stine, O.C., Morris, J.G. Jr.,
Sulakvelidze, A. and Sozhamannan, S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, The Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA
FEATURES
source 1..777
/organism="Vibrio cholerae"
/strain="571-88"
/db_xref="taxon:666"
/note="serogroup: O105"
1..777
/gene="ctxA"
1..777
/gene="ctxA"


```

55 AATGATGATAGTATATTCGGGCGAATTCAGACCTCCGTGATGAATAA 104
17 sargSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
105 GCATCTAGGTGCTCTTATGCCAAGAGCAGAAATGACTTCTTACCGAG 154
34 LyrHisMetAsnLeuAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
155 GACTCAAAATGATATCAACCTTATGATCATGCAAGAGCAATCAGACG 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
205 GGAATTTGTTAGCGAGATGATGATGTTCCACCTCAATTAAGTTGAG 254
67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTyrSerThrTyr 84
255 AAGTCCCACTTAGTGGGTCAAACTATATTCTCTGGTCATTTCTATT 304
84 YrlTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
305 ATATATATGTTATAGCCACCTGCACCAACATGTTTAACGTTAATGAT 354
101 LeuGlyValTyrSerProHisProTyrGlnGlnIleValSerAlaLeu 117
355 TTAGGGGATACAGTCTCATCAGATGACACAGAGTTTCTGCTTAGG 404
117 YGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnHeGly 134
405 TGGGATTCATACCTCCAAATATATGATGATGATCGAGTTTATTTGG 454
134 AlIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyr 150
455 TGCCTGATGACAACTATCATGTAATAGGGCTACAGAGATGATATTA 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
505 AGTACCTAGATGATGCTCCAGACAGATGTTATGATGAGATGAGTT 554
167 eProProAspHisGlnAlaTyrArgGluGluProTyrIleHisAlaP 184
555 CCCGCCGAGCATAGAGCTTGAGGAGAAAGCCGTGATTCATCATGAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleHrgIleAspThrCys 200
605 CGCGGGGTTGTGGGAATGCTCCAAAGATCATGATGAGTAATACGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrIleuArgGlyTyrGln 217
655 GAAAAAACCCAAAGCTAGTGTAAATCTCTGACGATCCAAATCTAA 704
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAlaPhe 234
705 AGTAAAGACAAATATTTTACGCTATCATCTGATATTTATACACATA 754
234 snArgIleArgAspGluLeu 240
755 ATAGATTAAGATGAATTA 774
seq_name: gb_ba:VIBCTXABA
seq_documentation_block:
LOCUS VIBCTXABA 1372 bp DNA linear BCT 01-FEB-2000
DEFINITION Vibrio cholerae ctxA gene and ctxB gene for cholera toxins,
complete cds.
ACCESSION D30052.1 GI:487332
VERSION 1
KEYWORDS ADP-ribosyltransferase; cholera toxin.
SOURCE Vibrio cholerae (sub-species O37, strain S7) (library: plasmid)
DNA, clone pKY310.
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 1372)
AUTHORS Yamamoto,K., Do,V.G., Xu,M., Iida,T., Miwatani,T., Albert,M.J. and

```

```

TITLE
Comparison of cholera toxin genes (ctxAB) of non-O1 vibrio cholerae
strains 854 (O139-bengal) and S7 (O37) from two outbreaks
unpublished (1994)
JOURNAL
REFERENCE
2 (bases 1 to 1372)
AUTHORS
Yamamoto,K.
TITLE
Direct Submission
COMMENT
Submitted (29-Apr-1994) Koichiro Yamamoto, Osaka University,
Institute for Microbial Diseases, Department of Bacterial
Infections: 3-1 Yamadaoka, Suita, Osaka 565, Japan
(Tel:06-879-4066)
Submitted (29-Apr-1994) to DDBJ by:
Koichiro Yamamoto
Institute for Microbial Diseases
Osaka University
3-1 Yamadaoka
Suita, Osaka 565
Japan
Phone: 06-879-4066
Fax: 06-879-8277.
FEATURES
source
1. 1372
/organism="Vibrio cholerae"
/strain="S7"
/sub_species="O37"
/db_xref="taxon:666"
/clone_lib="plasmid"
113..889
/gene="ctxA"
113..889
/gene="ctxA"
113..889
/gene="ctxA"
/codon_start=1
/product="cholera toxin"
/protein_id="BAA06288.1"
/db_xref="GI:808899"
/translation="WVKILFEFIFLSEFSYANDKLYRADSRPPEIKSGGLMPRG
ONEYEDRGTOMINILYDHARGTGTGFEVHDGYYVTSISLRSAILVGGTILSGSTYY
IYVITATNMENVNILGAYSPHDECEVSALGIPYSQIYGWYRVHFGVLDQLHRN
RGYRDYRYSNIDIAPADGYGAGPPEPRHMRPREPWTHNAPPCGNAPRSMNNTCD
EKTSGLGVKPLFDEYGVSKRQIFSGYSDIDTHRIKDEL"
113..166
/gene="ctxA"
886..1260
/gene="ctxB"
886..1260
/gene="ctxB"
/codon_start=1
/transl_table=11
/product="cholera toxin"
/protein_id="BAA06289.1"
/db_xref="GI:808900"
/translation="MIKIKFGVFYVLLSAYANGTPONTIDCAEYHNQIHLNDK
ILSYTESLAGNREMAIITFKGATFQVPEPSQIHDSKKAIERKKDRLRIAYILEAK
VEKICVWNKTPHAIALISMAN"
886..948
/gene="ctxB"
BASE COUNT 449 a 217 c 269 g 437 t
ORIGIN
sig_peptide
sig_peptide
BASE COUNT 449 a 217 c 269 g 437 t
ORIGIN
alignment_scores:
Quality: 1093.00 Length: 240
Ratio: 4.836 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 82.083
alignment_block:
US-09-528-682-1 x VIBCTXABA ..
Align seg 1/1 to: VIBCTXABA from: 1 to: 1372
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
|||||

```


167 AATGATGATAAGTATATATCGGACGATTCAGACCTCCTCGATGAATAA 216
 17 sargSerGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 217 GCAGTCAGGTGGCTTATGCGCAAGAGAGCAAGAGACTACTTGGACCGAG 266
 34 LyrHisMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 267 GTACTCAATGAATATCAACCTTATGATCATGCAAGACCAACCAACG 316
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 317 GGATTTGTATGGCAGCATGATGATGATGTTCCACCTCAATTAAGTTAG 366
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
 367 AAGGCCCACTTAGTGGTCAAACTATATGTCGTCATCTCACTTAT 416
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 417 ATATATATGTTATAGCCACTGCACCCCAACATGTTAACTTAATGATGTA 466
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
 467 TTAGGGGCATACAGTCTCATCCAGATGAACAAGACTTCTGCTTAGG 516
 117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
 517 TGGATTCCATACCTCCAAATATATGATGTTATCGATTCATTTTGGG 566
 134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 567 TGCCTGATGAACATTTACATCGTATATAGGGGCTACAGAGATAGATATAC 616
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 617 AGTAACTTATATATGCTCCAGCAGCATGTTATGATGTCGACAGGTT 666
 167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisHisAla 184
 667 CCCGCCGACATAGACCTTGGAGGAAGCCCTGATTCATCATCATCAC 716
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 717 CCGCGGCTTGGGATGCTCCAGATCATCATGATGATGATGATGATGAT 766
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
 767 GAAATAAACCAAGCTAGGTGTAATATCTTGCACCAATACCAATCTAA 816
 217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 817 AGTTAAAGACAAATATTTTCAGGCTATCAATCATGATATGATACATA 866
 234 snArgIleArgAspGluLeu 240
 867 ATAGATTTAAGATGAATTA 886
 seq_name: gb_pat:A16422

seq_documentation_block:
 LOCUS A16422 723 bp DNA linear PAT 06-JUL-1995
 DEFINITION A subunit of cholera toxin CT.
 ACCESSION A16422
 VERSION A16422.1 GI:641010
 KEYWORDS
 SOURCE
 ORGANISM
 . Vibrio cholerae.
 . Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 . IMMUNOGENIC DETOXIFIED MUTANTS OF CHOLERA TOXIN AND OF THE TOXIN
 . LT, THEIR PREPARATION AND THEIR USE FOR THE PREPARATION OF VACCINES

JOURNAL Patent: WO 9313202-A 45 08-JUL-1993;
 FEATURES Location/Qualifiers
 source 1..723
 /organism="Vibrio cholerae"
 /db_xref="taxon:666"
 1..723
 /gene="CT subunit A"
 <1..723
 /gene="CT subunit A"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA01281.1"
 /db_xref="GI:641011"
 /translation="NDKLYRADSRPDEIKOSGIMPRGSEYFDRGNMINLYDH
 ARGTOTGPRHDDGVSTISRSALHVGOTILSGSTYYIYVATAPRMKRVNVLG
 ATSPHDEQVSALGIPYSQLYGMTHFVGLDEQLHNRGTRDRTYSNLDIAPAD
 GYGLAFPEPRHMRREEPIWIIHAPPCGNAPRSSISNCTDEKTSGLGVFLDEYSKV
 KROIIFSGYSDIDTNRIRIDEL"

BASE COUNT 227 a 125 c 163 g 208 t
 ORIGIN

alignment_scores:
 Quality: 1088.00 Length: 240
 Ratio: 4.814 Gaps: 0
 Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
 US-09-528-682-1 x A16422 ..

Align seg 1/1 to: A16422 from: 1 to: 723

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIle 17
 117
 1 AATGATGATAAGTATATCGGCGAGATTCTAGACCTCCTCGTGAATAA 50
 17 sargSerGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 51 GCAGTCAGGTGGCTTATGCGCAAGAGAGCAAGAGACTACTTGGACCGAG 100
 34 LyrHisMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 101 GTACTCAATGAATATCAACCTTATGATCATGCAAGAGAACTACACGAG 150
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
 151 GGATTTGTATGGCAGCATGATGATGATGTTCCACCTCAATTAAGTTAG 200
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
 201 AAGTCCCACTTAGTGGGCTCAACTATATTTGTCGTCATCTTACTTAT 250
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 251 ATATATATGTTATAGCCACTGCACCCCAACATGTTAAGCTTAATGATGTA 300
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
 301 TTAGGGGCATACAGTCTCATCCAGATGACMAACAAGATTCTGCTTAGG 350
 117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
 351 TGGATTCCATACCTCCAAATATATGATGTTATCGATTCATTTTGGGG 400
 134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 401 TGCCTGATGAACATTTACATGCTAATAGGGGCTACAGAGATAGATATAC 450
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
 451 AGTAACTTATGATTTGCTCCAGCAGCATGTTATGATGATGTCGACGAG 500
 167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisHisAla 184

seq_name: gcb_pat:AR118598

LOCUS	723 bp	DNA	linear	PAT 16-MAY-2001
DEFINITION	Sequence 7 from patent US 6149919.			

VERSION AR118598.1 GI:14100508

SOURCE

Unclassified

AUTHORS Domenighini, M., Rapp

LT, their preparation and their use for the preparation of the final product.

FEATURES

Location/Qualifiers

```

PAGE COLUMN
337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 128
```

ORIGIN

```
alignment_scores:
  quality: 1088.00
  ratio: 4.814
  percent_similarity: 94.167
  length: 240
  gaps: 0
  percent_identity: 81.667
```

```
alignment_block:
```

Align seg 1/1 to: AR118598 from: 1 to: 723

[illegible]

```

251  AATATATATGTTATATACCCTGACCTGACCAACAGATGTTAAGTATGATGATGTA 3000
|||||
101  LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG1 117
|||||
301  TTAGGGGGGATACAGTCCATCCATCCAGAGTAAACAAAGAGTTCGTCTGTTAGG 3500
|||||
117  yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhleGly 134
|||||
351  TGGGATTCATACATCCCAAAATATATGATGATGATATCAGATTCATTTGCGG 4000
|||||
134  allLeaspGluIarGluLeuHisArgAsnArgIuTyrArgAspArgTyrTyr 1500
|||||
401  TCGTTGATGAAACAATTACATCGTATATAGGGGCTTACACAGATACATATTAC 4500
|||||
151  ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||
451  AGTAACTTAGATATTGCTCCACAGCAGATGCTATGATTCGACAGGTTT 5000
|||||
167  eProProAspHisGlnAlaIATrPArgGluGluProTribLeHisHisAlaP 184
|||||
501  CCCTCGGGAGCATATAGCTTGAGAGGAAGACCGCTGATTCATCATCAGCAC 5500
|||||
184  roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
551  CCGCCGGGTGGGGAGATCTCCACAGATCATCGATCACTAATACTTCGAT 6000
|||||
201  GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIy 217
|||||
601  GAAAAAACCCAAAGCTAGGTGAATAATTCCTTAGACATATACCATCTAA 6500
|||||
217  sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
|||||
651  AGTTAAAGACAAATATTTTCAGCTCATCATCGATATGATATGATACACATA 7000
|||||
234  snArgIleArgAspLeuIeu 240
|||||
701  ATAGATATTAAGGATGATTA 720

```

seq_name: gfb_pat:E00132

```
seq_documentation_block:
```

DEFINITION	DNA coding of cholera toxin.
100133	

VERSION E00132.1 GI:2168431

SOURCE unidentified.

unclassified.

AUTHORS Arufuooru, N. and Do, B.M. :
EDITOR CRIVELLO, A. AND P. CE DUA ADVANCEMENT RECOMBINED DNA CHOICEA TOYIN

TOURNAMENT AND MEDICINE
 Boston, TD 1003333033-A 1 33-DEC-10033

COMMENT

PI ARUFUOORU NIJIERU, DO BUIRUDO MISHIERU
PC A61K39/106 A61K39/00 C07H21/04 C13N15/00-9

alignment_scores:

Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:

US-09-528-682-1 x VCTOXAB21 ..

Align seg 1/1 to: VCTOXAB21 from: 1 to: 1362

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
223 AAGGATGATAGTATATCGGCGAGATTCAGACCTCCATGATGATAAAAA 272
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
273 GCAGTCAGGIGGCTTATGCCAGAGGACAGAGTACTTTGACCGAG 322
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
323 GTCATCAATGAAATATCAACCTTATGATCATGCAGAGAGGACTCAGACG 372
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 GlyPheAlaArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
373 GGATTTGTTAGGACGATGATGATATGTTCCACCTCAATAGTTGAG 422
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
423 AAGTGCACCATTTAGTGCGTCAACATATTTCTGCTGCTACTTACTTAT 472
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84 yrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
473 ATATATATGTTATAGCCATGCACCAACATGTTTAACTTAATGATGATGA 522
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 LeuGlyValTyrSerProHisProTyrGluGluGluValSerAlaLeuGl 117
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
523 TTVGGGGCCATACAGTCTCCATCCAGATGAAAGAAAGTTCTCTCTTAA 572
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
573 TGGGATTCCTACTCCCAATATATGATGATGATCCAGTTCTTTGGGG 622
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
623 TCGTGTGATGAACAATATACATCGTAATAGGGGCTACAGAGATATATAC 672
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgGluAlaGlyPh 167
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
673 AGTAACTTATGATATGCTCCAGCAGCAGATGATGATGATGATGATGAT 722
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 eProProAspHisGlnAlaTyrArgGluGluProTyrPheHisAlaIle 184
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
723 CCTCCGAGACATAGAGCTTGGAGGAGAGCGGTGATTCATCATGCGAC 772
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184 roGlnIleCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
773 CGCCGGGTGCTGGGAATGCTCCAGATCATCATGATGATGATGATGAT 822
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIy 217
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
823 GAAAGAAACCAAGCTCAGGTGAAATTCCTTGCAGAAATCAACCAATCTAA 872
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
873 AGTTAAAGCAATATTTTCAGGCTATCATCTGATATTTGATACACATA 922
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234 snArgIleArgAspGluLeu 240
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
923 ATGAAATTAAGATGAATTA 942
  ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: gb_ba:VIBCTXAB

seq_documentation_block:

```

LOCUS          VIBCTXAB                      1369 bp    DNA        linear    BCT 01-FEB-2000
DEFINITION     Vibrio cholerae ctxA gene and ctxB gene for cholera toxins,
complete cds.
ACCESSION      D30053
VERSION        D30053.1 GI:487333
KEYWORDS       ADP-ribosyltransferase; cholera toxin.
SOURCE         Vibrio cholerae (sub-species O139-Bengal, strain 1854) (library:
plasmid) DNA, clone PKY340.
ORGANISM       Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision: Vibrionaceae; Vibrio.
1 (bases 1 to 1369)
Yamamoto, K., Do, Y.G., Xu, M., Iida, T., Miwatani, T., Albert, M.J. and
Honda, T.
Comparison of cholera toxin genes (ctxAB) of non-O1 Vibrio cholerae
strains 854 (O139-bengal) and S7 (O37) from two outbreaks
unpublished (1994)
2 (bases 1 to 1369)
Yamamoto, K.
Direct Submission
Submitted (29-APR-1994) Koichihiro Yamamoto, Osaka University,
Institute for Microbial Diseases, Department of Bacterial
Infections; 3-1 Yamadaoka, Suita, Osaka 565, Japan
(Tel:06-879-4066)
Submitted (29-Apr-1994) to DDBJ by:
Koichihiro Yamamoto
Institute for Microbial Diseases
Osaka University
3-1 Yamadaoka
Suita, Osaka 565
Japan
Phone: 06-879-4066
Fax: 06-879-8277.

FEATURES             source          location/Qualifiers
     source          1..1369
     /organism="Vibrio cholerae"
     /strain="1854"
     /sub_species="O139-Bengal"
     /db_xref="taxon:666"
     /clone_lib="plasmid"
     110..886
     /gene="ctxA"
     110..886
     /gene="ctxA"
     /codon_start=1
     /transl_table=1
     /product="cholera toxin"
     /protein_id="BA06290.1"
     /db_xref="GI:808901"
     /translation="MWKIIFFVFIFLSSFSYANDKLYRADSRPDELKQSGIMPRG
OSEYEDRGTQNMNIMLYDHARCTGFVPHDGVYSTISLSAHLYVQGTILSGSTYY
IYVATAPNMENVDVIGAYSPHDEDEVSALGIPYSQIGVRFHGVLEDELHNRN
RGYDARYSNLDIAPADGVGLAFPPPHRAMREPMTHHAPPCGNAFRSMNTCD
EKQSLGCKFLDEYQSKYKRIFSGYQSDIDTNRKIDEL"
     110..163
     /gene="ctxA"
     883..1257
     /gene="ctxB"
     883..1257
     /gene="ctxB"
     /codon_start=1
     /transl_table=1
     /product="cholera toxin"
     /protein_id="BA06291.1"
     /db_xref="GI:808902"
     /translation="MWIKRGEVFTVILSSYANGTPOINTDCAEYHNRQIYTLNDK
IEFSTESLAGKREMAIITFKNGALFQVEVPSQIHIDSKAIERMKDPLRIAYLLEAK
VEKLCVNNKTPHAIASMAN"
     883..945
     /gene="ctxB"

sig_peptide
sig_peptide
BASE COUNT      449 a      214 c      269 g      437 t
ORIGIN

```


alignment_scores:

Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:

US-09-528-682-1 x VIBCTXABB ..

Align seg 1/1 to: VIBCTXABB from: 1 to: 1369

```

1 AsnGLYAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
164 AATGATGATTAAGTATTAATCGGGCAGATTTAGACCTCCGATTAATATA 213
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
214 GCAGTCAGGTGGCTTATGCCAAGAGACAGAGTACTTACTTGACCGAG 263
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
264 GATCTCAAAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGACG 313
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
314 GGATTTGTTAGCAGCATGATGATGTTCCACCTCAATTAGTTTGAG 363
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
364 AAGTCCCACTAGTGGTCAAACTATATGTCGTGCTCATCTTCACTTAAT 413
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84 yIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
414 ATATATATGTATNAGCCACTGCACCCACATGTTTAACGTTAATGATGA 463
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
464 TTAGGGGCGATACAGTCTCATCCAGATGAACAAGAGTTTCTGCTTAGG 513
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPtyrArgValAsnPheGly 134
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
514 TGGGATTCATCTACTCCCAATATATGATGATGTCAGTTCAATTTGGGG 563
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
564 TGGTATGATGACATATACATCGTAATAGGGGCTACAGAGATGATATTAC 613
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
614 AGTAAGTATGATATTTGCTCCAGACAGATGTTATGATTTGGCAGATT 663
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 eProProAspHisGlnAlaIleTyrArgGluGluProTyrIleHisAlaP 184
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
664 CCTCCGCGAGCATAGAGCTTGGAGGAAAGCCGCGATTCATCATGACG 713
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184 rGlnGlyGlyGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
714 CGCCGGGTTGTGGCAATGCTCCAGATCATCGATGAGTAACTATTCGAT 763
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
764 GAAAGAACCCAAAGTCTAGTGTAAATTCCTTGACGAATACCAATCTAA 813
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
814 AGTTAAAGAACAAATATTTTCAAGGCTATCATCTGATTAATGATACACA 863
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234 snArgIleArgAspGluLeu 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
864 ATAGAAATTAAAGATGAATTA 883
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: gb_ba:VCTOXA56

seq_documentation_block:

LOCUS VCTOXA56 1397 bp DNA linear BCT 30-JUN-1993
DEFINITION V.cholerae genes for toxin proteins A and B, strain 569B.
ACCESSION X58785 S53782
VERSION X58785.1 GI:48888
KEYWORDS cholera toxin.
SOURCE Vibrio cholerae.
ORGANISM Vibrio cholerae.
Bacteria: Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 1397)
AUTHORS Dams,E.
TITLE Direct Submission
JOURNALS Submitted (25-MAR-1991) E. Dams, Laboratory of Human Biochemistry,
RUGA University Antwerp, Gruenembergerlaan 171, B-2020 Antwerp,
Belgium
2 (bases 1 to 1397)
REFERENCE Dams,E., De Wolf,M. and Dierick,W.
TITLE Nucleotide sequence analysis of the CT operon of the Vibrio
JOURNALS cholerae classical strain 569B
MEDLINE Biochim. Biophys. Acta 1090 (1), 139-141 (1991)
91355224
FEATURES
source
Location/Qualifiers
1..1397
/organism="Vibrio cholerae"
/strain="classical strain 569B"
/db_xref="taxon:666"
/clone="PUM17"
72..127
/rpt_type="RANDM
145..156
171..177
204..980
/codon_start=1
/transl_table=1
/product="cholera toxin A protein (CTA)"
/protein_id="CAA41590.1"
/db_xref="GI:48889"
/db_xref="SWISS-PROT:P01555"
/translation="MVKITVEFTLSEFSYANDDKLYRADSRPPDEIKQSGGLMPRG
QSEYEDRGTONNINLYDHARGTGTGFEVHDGIVSTISLSAHLYGOTILSGSTTY
IYVATAVNMFENVLDGAVSPHPDDEVSALGIPYQIYGVYRVHFGVLDQHLRN
RGYDRYASNDIADPADGYGLAGPPREHRAHREPPIHNAAPCGGNAPRSSMSWTC
EKTSIGYKFLDEYQSKVKRQIFSGYQSDIDTHNRINDEL"
204..257
258..977
977..1039
/product="cholera toxin A protein (CTA)"
977..1039
/codon_start=1
/transl_table=1
/product="cholera toxin B protein (CTB)"
/protein_id="CAA41591.1"
/db_xref="GI:48890"
/db_xref="SPTREMBL:057193"
/translation="MIKKRGVFTVLLSSAYANGTPONTIDCAEYHNRQIHLNDK
IFSTTESLAGREMAITFKNGATFQVEYPSQIHDSKAIENKDTLRATLYLEAK
VEKICWNKRPDAIATSMAN"
1040..1348
/product="cholera toxin B protein (CTB)"
BASE COUNT 460 a 207 c 281 g 449 t
ORIGIN

alignment_scores:

Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:

US-09-528-682-1 x VCTOXA56 ..

Align seg 1/1 to: VCTOXA56 from: 1 to: 1397

1 AsnGIyspArgLeuTYrArgAlaIspSerArgProProAspGluIleIy 17
238 AATGATGATGAAGTTATATCGGCGAGAGTCTTACACCTCTGATGAATAATA 307
17 sArgSerGIyGIyLeuMetProArgGIyHISasnGIyTYrPheAspArg 34
308 GCAGTCAGAGTGTCTTATGTCACAAAGACAGAGTGAGTACTTTCAGCCGAG 357
34 IYThGIImetAsnIleasnLeuTYrAspHisAlaArgGIyThrGlnThr 50
358 GTACACCAATGAAATATCAACCTTTATGATCAAGCAAGAGAACTCGAGAG 407
51 GIlPheValArgTYrAspAspGIyTYrValSerThrSerLeuSerLeuArg 67
408 GGATTTGTTAGGACGATGATGATGATGTTTCCACCTCATTAAGTTTGAG 457
67 gSerAlaHisIleuAlaGIyGlnSerIleLeuSerGIyTYrSerThrTYrT 84
458 AAGTGGCCACTTAGTGCGTCACAACTATATGCTGTGCTACTTCTACTATAT 507
84 YrIleTYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
508 AATATATATGTATTAGCCACTGCACCCACCAATGTTTAACTTATATGATGA 557
101 LeuGIyValTYrSerProHisProTYrGIyGlnGlnValSerAlaIleuGI 117
558 TTAGGGGCGCTACAGTCCCTCATCCAGATGACAAAGAGTTTCGCTTTAGG 607
117 yGIyIleProTYrSerGlnIleTYrGIyTYrPyrArgValAsnPheGIyV 134
608 TGGGATTCATCTACCTCCCAATATATGATGATGATGATGATCTATTTTGGGG 657
134 aIleAspGIuArgLeuHisArgAsnArgGIuTYrArgAspArgTYrTYr 150
658 TCGTGTGATGAACAATTACATCGTAATAGGCGCTACAGAGATGATGATATAC 707
151 ArgAsnLeuAsnIleAlaProAlaGIuArgGIyTYrArgLeuAlaGIyTh 167
708 AGTAACTTGAATATGCTCTCCACGACGAGATGGTTATGATTTGGCAGAGTT 757
167 eProProAspHisGlnAlaIATPrArgGIuArgProTYrIleHisAlaIAP 184
758 CCCTCCGAGCATAGAGCTTGGAGGAGAAAGCCCGTGGATTCATCATGCAAC 807
184 rGlnGlnIyCysGIyAsnSerSerArgThIleThrGIyAspThrCysAsn 200
808 CGCCGGGTGTGGGAATGCTCCACAGATCATGATGATGATGATGATGATGAT 857
201 GIuGIuThrGlnAsnLeuSerThrIleTYrLeuArgGIuTYrGlnSerIy 217
858 GAAAAAACCCAAAGTGTAGTGTAAATATCTTGGACGATATACCAATCTTA 907
217 sValIyAsArgGlnIlePheSerAspTYrGlnSerGIuValAspIleTYrA 234
908 AGTTAAATAACAATATTTTTCAGCGCTATATCTGATATGATATACACATA 957
234 snArgIleArgAspGIuLeu 240
958 ATAGAAATTAAGATGATTA 977
seq_name: gb_ba:VCCITX
seq_documentation_block:
LOCUS VCCITX 2020 bp DNA linear BCT 30-MAR-1995
DEFINITION Vibrio cholera toxin (ctx) operon DNA sequence from strain 2125.
ACCESSION X00171
VERSION X00171.1 GI:48347
KEYWORDS overlapping genes; signal peptide; tandem repeat.
SOURCE Vibrio cholerae.
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
1 (bases 1 to 2020)
AUTHORS Mekalanos,J.J., Swartz,D.J., Pearson,G.D., Harford,N., Groyne,F.

[illegible]

570 AATGATGATAGTATATATCGGCGAGATTCCTAGACCTCTGATGAATAA 619
17 sargSerGlyLeuMetProArgGlyHisnGlyIleuArgPheAspArg 34
620 GCAGTCAGGTCGTATATGCAAGAGACAGATGATCTTGGACCGAG 669
34 lYthGlnMetAsnIleAsnLeuTyArgPheHisAlaArgGlyThrGlnThr 50
670 GTACCTCAATGATATCAACCTTTATGATCATGCAAGAGAACTCAGACG 719
51 GlyPheValArgTyArgAspArgGlyTyArgValSerThrSerLeuSerLeuArg 67
720 GGATTTGTAGCGACAGATGATGATATGTTTCCACCTCAATAGTTTGGAG 769
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyArgSerThrTyT 84
770 AAGTCCCGCCTTAGTGGTCAACATATATGCTGCTGCTCTTACTATAT 819
84 yrlIeTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
820 ATATATATGTTATAGCCACTGACCCACATGTTTAACTTATATGATGTA 869
101 LeuGlyValTyArgSerProHisProTyArgGlnGluValSerAlaLeuG 117
870 TTAGGGGCATACAGTCCTCATCCAGATGAAACAAGATTCTGCTTAGG 919
117 yGlyIleProTySerGlnIleTyArgTyArgValAsnPheGly 134
920 TGGGATTCATATCTCCCAATATATGATGATGATGATGATGATGATG 969
134 alIleAspGluArgLeuHisArgAsnArgGlyTyArgAspArgTyTyr 150
970 TCCCTGATGAACAATATATCTGTAATAGGGCTACAGACATATGATATTC 1019
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyArgLeuAlaGlyP 167
1020 AGTAACTGATGATATGCTCCACAGACAGATGCTTATGATGCGAGTTT 1069
167 eProProAspHisGlnAlaThrArgGlnGluProTrpIleHisAlaP 184
1070 CCTCGGAGCATATAGCTTGGAGGAGAGCCGCGATTCATCATGCGAC 1119
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
1120 CCGCGGGTGTGGGATGCTCCAGATCATGATCATGATCATGATCATG 1169
201 GluGluThrGlnAsnLeuSerThrIleTyArgLeuArgGlyTyArgSer 217
1170 GAAAAAACCCAAAGCTTAGGCTTAAATTCCTTGACGATACCAATCTTA 1219
217 sValIysArgGlnIlePheSerAspTyArgIleSerGluValAspIleTy 234
1220 ACTTAAAGACAAATATTTTTCAGGCTATCATATGATGATGATGATG 1269
234 snArgIleArgAspGluLeu 240
1270 ATAGAATTAAAGATGATTA 1289

seq_name: gb_pat:E03512
seq_documentation_block:
LOCUS E03512 2020 bp DNA linear PAT 29-SEP-1997
DEFINITION Cholera toxin gene.
ACCESSION E03512
VERSION E03512.1 GI:2171728
KEYWORDS JP 1992099488-A/17.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2020)
AUTHORS Shirai,H.
TITLE PRIMER FOR AMPLIFICATION OF GENE
JOURNAL Patent: JP 1992099488-A 17 31-MAR-1992;

UNITIKA LTD
PN JP 1992099488-A/17
PD 31-MAR-1992
PF 20-AUG-1990 JP 1990219523
PI SHIRAI HIROMASA
PC C12N15/11.C1201/68;
CC strandedness: Double;
FH topology: Linear;
FH key Location/Qualifiers
FT misc_feature 1..2020
FT Location/Qualifiers
FEATURES
source 1..2020
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 602 a 352 c 446 g 620 t
ORIGIN
alignment_scores:
Quality: 1086.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667
alignment_block:
US-09-528-682-1 x E03512 ..
Align seg 1/1 to: E03512 from: 1 to: 2020
1 AsnGlyAspArgLeuTyArgAlaAspSerArgProProAspGluIleTy 17
1170 GCGAGTCAGGTCGTATATGCAAGAGACAGATGATGATGATGATGATG 669
570 AATGATGATAGTATATATCGGCGAGATTCCTAGACCTCTGATGAATAA 619
17 sargSerGlyLeuMetProArgGlyHisnGlyIleuArgPheAspArg 34
620 GCAGTCAGGTCGTATATGCAAGAGACAGATGATGATGATGATGATGATG 669
34 lYthGlnMetAsnIleAsnLeuTyArgPheHisAlaArgGlyThrGlnThr 50
670 GTACCTCAATGATATCAACCTTTATGATCATGCAAGAGAACTCAGACG 719
51 GlyPheValArgTyArgAspArgGlyTyArgValSerThrSerLeuSerLeuArg 67
720 GGATTTGTAGCGACAGATGATGATATGTTTCCACCTCAATAGTTTGGAG 769
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyArgSerThrTyT 84
770 AAGTCCCGCCTTAGTGGTCAACATATATGCTGCTGCTCTTACTATAT 819
84 yrlIeTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
820 ATATATATGTTATAGCCACTGACCCACATGTTTAACTTATATGATGTA 869
101 LeuGlyValTyArgSerProHisProTyArgGlnGluValSerAlaLeuG 117
920 TGGGATTCATATCTCCCAATATATGATGATGATGATGATGATGATG 969
134 alIleAspGluArgLeuHisArgAsnArgGlyTyArgAspArgTyTyr 150
970 TCCCTGATGAACAATATCTGTAATAGGGCTACAGACATATGATGATG 1019
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyArgLeuAlaGlyP 167
1020 AGTAACTGATGATATGCTCCACAGACAGATGCTTATGATGATGATG 1069
167 eProProAspHisGlnAlaThrArgGlnGluProTrpIleHisAlaP 184
1070 CCTCGGAGCATATAGCTTGGAGGAGAGCCGCGATTCATCATGCGAC 1119


```

2276 TGGGATTCATACCTCCCAATATATGATGATGATTCATTTGGG 2325
134 allieaargluargleuHisargasnargglutylararaspargtyrtyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2226 TGGTGTGATGAACCAATATCATCTGATATAGGGGCTACAGAGATGATATTAC 2375
151 ArgasnleuasnilleaIleProAlaGluaspGlytyrtyrargleuAlaGly 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2376 AGTAACTTAGATATTGCTCCAGCAGCAGATGCTATGATGGCAGGTT 2425
167 eProCaspHisGlnAlaIleTparargGluGluProTyrPheHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2426 CCGTCCGAGACATGACCTTGAGGAGAACCCCTGATTCATCATGCAC 2475
184 roGInGlyCysGlyAsnSerSerargThrIleThrGlyAspThrCysasn 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2476 CGCCGGGTGGGAAATGCTCCAGATCATGATGAGTAATCATCTGCAT 2525
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2526 GAAAAAACCCAAAGCTAGTGTAAATTCCTTGACGATACCAATCTAA 2575
217 SValIysargGlnIlePheSeraspTyrGlnSerGluValaspIleTyr 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2576 AGTTAAAGACAAATATTTTCAGGCTATCATCTGATATGATACACATA 2625
234 snArgIleArgaspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2626 ATGAGATTAAAGATGAATTA 2645

seq_name: gb_AF414369

seg documentation block:
LOCUS AF414369 4275 bp DNA linear PIC 02-OCT-2001
DEFINITION Vibrio phage CTX Ace (ace), Zot (zot), CtxA (ctxA), and CtxB (ctxB)
ACCESSION AF414369
VERSION AF414369.1 GI:15825396
KEYWORDS
SOURCE
ORGANISM
Vibrio phage CTX.
Virus; ssDNA viruses; Inoviridae; Inovirus.
REFERENCE
1 (bases 1 to 4275)
AUTHORS Bhattacharyya,T., Nandy,R.K. and Nair,G.B.
TITLE The entire core region of the ctx-phi (ctx-phage) in VCE 232, an
environmental strain of V. cholerae
JOURNAL Unpublished
2 (bases 1 to 4275)
AUTHORS Bhattacharyya,T., Nandy,R.K. and Nair,G.B.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2001) Microbiology, NICED, Calcutta, India
FEATURES
source
1..4275
/organism="Vibrio phage CTX"
/specific_host="Vibrio cholerae strain VCE 232"
/db_xref="taxon:141904"
/note="prophage: core region"
1543..1833
/gene="ace"
1543..1833
/gene="ace"
1543..1833
/gene="ace"
/codon_start=1
/transl_table=11
/product="Ace"
/protein_id="AAL09683.1"
/db_xref="GI:15825396"
/translation="MLMDTLIDYDMLIDGFTWLVIKGIIMIESKIFVIOFEMWSQKV
IDMTIYPLIQADMLDPQYSGFLFFGLDQALAIIVLQALMTFRALRLNT."
1830..3029
/gene="zot"
1830..3029
/gene="zot"
/codon_start=1

```

```

/transl_table=11
/product="zot"
/protein_id="AAL09684.1"
/db_xref="GI:15825400"
/translation="MSIFLHAGPQSYKTSGLALMLRLPAIKSGRIITNVKGLNER
MAKYLKMDVSDISIEFIDTHDGRITLRFNMMAKDAFLFIDECGRIMPRLTATN
LKALDPDLVADRESEVAFDMRRHGMQICLTPPIAKVHNHRAAEIGYHF
NRATYGLAKFLLTHDANSGOMSHALTROVKRIPSPIFKMYASTTGKARDTMAG
TALMDRKTLFLFGWFLMFESYSPFGLNHPFTFGNDMTTISEOSEPSKATAGAV
GSAVAPASFEFCIGRLVQDFVTVDERTYLVNDLDPYGLMATGHNHITKDLTV
FEETSGSVTEFLFASVRYKVLPLDFNHVVFDFFAQALWVERRLPIKTENDK
KGLNSIF"
3128..3904
/gene="ctxA"
3128..3904
/gene="ctxA"
3128..3904
/gene="ctxA"
/codon_start=1
/transl_table=11
/product="CtxA"
/protein_id="AAL09681.1"
/db_xref="GI:15825397"
/translation="MYKILFVFIFLISFSYANDKLYRADSRPDEIKOSGLMPRG
OSEYEDRGQNMNINLYDHRGTQTFVRHDDGVSTISLRSANLVGOTILSGSTYY
IVIVATAPMFMVNDVLAISPDEQVSAAGIPIYSQIYGVRYHFGVLDLHRN
RGYRDYVNLDIAPADGYGLAGFPPEBAREBEMIHHPGCGNAPRSSMSNTCD
EKTOSIGVKEFLDEYOSKVRQIFSGYQSDIDHNRIKDEL"
3901..4275
/gene="ctxB"
3901..4275
/gene="ctxB"
/codon_start=1
/transl_table=11
/product="CtxB"
/protein_id="AAL09682.1"
/db_xref="GI:15825398"
/translation="MIKIKFGVFTFVLLSSAYAGHPHNITLCAEYHNTQIHTLNDK
IFSYSFLAGKREMAITIFKNGATFOVEVPGSHQIDSKAIERMDQFLRIYLVLEAK
VERICWNKKTPHAIATISMAN"

BASE COUNT 1132 a 904 c 997 g 1242 t
ORIGIN

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
us-09-528-682-1 x AF414369 ..
Align seg 1/1 to: AF414369 from: 1 to: 4275

1 AsnGlyaspArgLeuTyrArgAlaaspSerargProProaspGluIlely 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3182 AATGATGATTAGTATATCGGCGAGATTCTGACCTCTGATGAATAAA 3231
17 sarGserGlyGlyLeuMetProArgGlyHisAsnGluTyrPheaspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3332 GCACTCAGGTGGTCTTATGCCAAGGACAGAGACTTGTACCGAG 3281
34 lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3282 GTATCTCAATGAATATCACTTTATGATCTCGCAAGAGACTCAGACG 3331
51 GlyPheValArgTyrAspaspGlyTyrValSerThrSerLeuSerLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3332 GGATTTGTTAGCAGCATGATGATGTTCCACACCAATATAGTTGAG 3381
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3382 AAGTCCCACTTACTGCTCAACTATATGCTGTCATCTACTTATTT 3431
84 yrlIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnaspVal 100

```


3432 ATATATATGTATAGCGACTGCACCCACCATGTTTAACTTATATGATGA 3481
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG1 117
3482 TTATGGGGCGATACAGTCCATCCATCCAGATGAACAAGAAGTTTCCTTTAGG 3531
117 yGlyIleProTyrSerGlnIleTyrGlyIleTyrPheValAsnPheGlyV 134
3532 TGGGATTCATCTACCTCCCAATATATGATGCTGATTCAGATTCATTTGGGG 3581
134 a11IleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
3582 TCCTTGATGACACATATACATCGTAATAGGGCGCTACAGATATGATATAC 3631
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaIleYh 167
3632 AGTAACCTTGATATGCTCCAGCAGCAGATGGTATGATATGGCAGGTTT 3681
167 eProPAspHisGlnAlaTPrArgGluGluProTPrIleHisAlaP 184
3682 CCTCCGGAGCATAGAGCTTGGAGGGAACGCCGTGATTCATCATGCAC 3731
184 roGlnIleCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
3732 CCGCGGGTTGTGGGAATGCTCCAGATCATCATGATGATATCTTCGAT 3781
201 GluGluThrGlnAsnLeuSerThrIleTyrIleArgGluTyrGlnSerly 217
3782 GAAAAACCCAAAGTACGTGATGTAATTCCTTGACGAATACCAATCTAA 3831
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleYh 234
3832 AGTTAAACACAAATATTTCAGGCTATCATCTGATATGATACACATA 3881
234 snArG1IleArgAspGluLeu 240
3882 ATRGAAATTAGAGTGAATTA 3901

seq_name: gb_ba:AE004224

seq_documentation_block:
LOCUS AE004224 10646 bp DNA linear BCT 31-JUN-20000
DEFINITION *Vibrio cholerae* chromosome I, section 132 of 251 of the complete
chromosome.
ACCESSION AE004224 AE003852
VERSION AE004224.1 GI:9655952
KEYWORDS
SOURCE
ORGANISM *Vibrio cholerae*.
vibrio cholerae
Bacteria: Proteobacteria; gamma subdivision: Vibrionaceae: *Vibrio*.
REFERENCE
1 (bases 1 to 10646)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermolova, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nieman, W.C. and White, O.
DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
cholerae
Nature 406 (6795), 477-483 (2000)
2 (bases 1 to 10646)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermolova, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nieman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Metkunas, J.J., Venter, J.C. and Fraser, C.M.
Direct Submission
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA

FEATURES	source	location/Qualifiers
gene	CDS	1..10646 /organism="Vibrio cholerae" /strain="N16961" /serotype="O1" /db_xref="taxon:666" /chromosome="1" /note="biotype: El Tor" complement(215..439) /gene="VC1452" complement(215..439) /gene="VC1452" complement(215..439) /note="similar to GP:2564354; identified by sequence similarity: putative" /codon_start=1 /transl_table=11 /product="RstC protein" /protein_id="AAF94609.1" /db_xref="GI:9655953" /translation="MSLKPYTLMDVDSLELDNNMALYRSGAVTDEIAHQVNLICD KIIDGIVNFRIRSPINPOLKATPEPSL" complement(533..907) /gene="VC1453" complement(533..907) /gene="VC1453" /note="similar to GP:2564353; identified by sequence similarity: putative" /codon_start=1 /transl_table=11 /product="RstH protein" /protein_id="AAF94610.1" /db_xref="GI:9655954" /translation="MMKLIYINMKSRVVEGASHSEGVSTGAPYLLPVLEVGKPIRQ WKNKGGCLTFGLQHOEVKFEVSDPAMTRKLEQAFVLYTFDNEPDEPSRLVIDY OVVSLEFDNPGGKPLDKPNKI" complement(885..1964) /gene="VC1454" complement(885..1964) /gene="VC1454" /note="similar to GP:2564352; identified by sequence similarity: putative" /codon_start=1 /transl_table=11 /product="RstI protein" /protein_id="AAF94611.1" /db_xref="GI:9655955" /translation="MKKQIFTLDELQDITNASPEFVYDLAMSVPYASFRAHKSDLS SLIAPLPKPDYRMARPEQEKLEIYKOKWYAMMERLEVEICLHYLGIRMSPWDRK GLYVENSCHSLKYSNKHVGFANLGGNRTCYFOIEGVCRTYLEHTSLPRLMWD ILGCSRLTIDAVDDPHGLPGREYARKAVSDAEPFARAPANGGERLYSPNGKI IINSEFVGSRESRIYMRITYNKAOLGIDMHMFRRREYLRKMPIDVILNITGYPAGICA YSASITNSLPKVVTKRQVALDHSHIKWARRVGKTLFDISKHFGDLERVFAGLI SKELHSDSLNLPDSYMKLIDEIMGD" 2090..2428 /gene="VC1455" 2090..2428 /gene="VC1455" /note="similar to GP:2564351; identified by sequence similarity: putative" /codon_start=1 /transl_table=11 /product="transcriptional repressor RstR" /protein_id="AAF94612.1" /db_xref="GI:9655956" /translation="MKIKERLANORKIINKTQAMADEIGISLSTYKKYSGSGLPTM ENLYKDALEISIDELGCRWATDENGELMLRLKKIQDLDEDRKATSWLESGLINH STKSLNHGA" complement(3030..3404) /gene="VC1456" complement(3030..3404) /gene="VC1456" /note="similar to GB:K01170 SP:P01556 GB:D30052 GB:D30053 GB:K58785; identified by sequence similarity: putative"


```

1 AsnGlyAspArgLeuIyrArnAlaAspSerArgProProAspGluIleIy 17
11 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
196 AARGATGATGATGATTATATCGGCGAGATTGTAGACCTCGATGAAATAA 245
17 sArgSerGIyGIyLeuMetProArgGIyLysAnGluIyTyRpheAspArg 34
34 |yThGIlnMeLAsnIleAsnLeuTyRAspHisAlaArgGIyThrGlnTr 50
296 GFACCTAAATGATATCAACCTTTATGATCATGCAAGAGAACTCAGAG 345
51 GlyPheValArgTyRAspAspGIyTyRValSerThrSerLeuSerLeu 67
346 GGAATTGTTAGGCGACGATGATGGATATGTTTCCACCTCATTTAGT 395
67 gSerAlaHisLeuAlaGIyGlnSerIleLeuSerGIyTyRSerThrTy 84
396 AAGTGGCCACTTAGTGTGGTCAACATATTTGCTGTACTTACTTATT 445
84 yTlleTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
446 ATATATATGTTATAGCCACCTGCACCCCAACATGTTTAACTTATGAT 495
101 leuGIyValTyRSerProHisProIyTyRGIyGlnGluValSerAla 117
496 TTGAGGGCATACAGTCTCTCATCCAGATGACANAGAGATTTCGCTTA 545
117 yGIyIleProTyRSerGlnIleTyRGIyTyRPTyRArgValAsnPh 134
546 TTGGATTTCATACTCCCAATATATGATGATGATGATGATGATGATG 595
134 alIleAspGIyArgLeuHisArgAsnArgGIyTyRArgAspArgTy 150
596 TGGTGTATGATGACATTTACATCGTAATAGGGCGCTACGAGATGAT 645
151 ArgAsnLeuAsnIleAlaProAlaGluAspGIyTyRArgLeuAla 167
646 AGTAACTTATGATTTGCTCCAGCAGCAGCATGTTATGATTTGGCAG 695
167 eProProAspHisGlnAlaTrpArgGluGluProTrpIleHisHisAl 184
696 CCTCGGAGCATATAGCTTGGAGGAGAAAGCCGTTGATTCATCATG 745
184 roGlnGIyCysGlyAsnSerSerArgThr 193
746 CGCGGGGTGTGGGAATGCTCCAGAGATCA 774

seq_name: gb_ba:ECOETOXHL

seq_documentation_block:
LOCUS      ECOETOXHL               1262 bp    DNA       linear    BCT 18-NOV-1994
DEFINITION Escherichia coli (strain 41) heat-labile enterotoxin type IIB
            (LT-IIB) A and B chain genes, complete cds.
ACCESSION  M28523
VERSION    M28523.1 GI:576584
KEYWORDS   LT-IIB gene; enterotoxin type IIB.
SOURCE     Escherichia coli (strain 41) DNA.
ORGANISM   Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE  1 (bases 1 to 1262)
            Pickett, C.L., Twiddy, E.M., Coker, C. and Holmes, R.K.
            Cloning, nucleotide sequence, and hybridization studies of the type
            IIB heat-labile enterotoxin gene of Escherichia coli
            J. Bacteriol. 171 (9), 4945-4952 (1989)
JOURNAL    MEDLINE
COMMENT    On Nov 28, 1994 this sequence version replaced gi:341953.
FEATURES
            location/Qualifiers
                1..1262
                /organism="Escherichia coli:"
                /strain="41"

```



```

/db_xref="taxon:562"
/clone="pcc4185"
-35_signal
1..12
/gene="LT-IIB"
1..1262
/gene="LT-IIB"
26..31
/gene="LT-IIB"
113..904
/gene="LT-IIB"
/note="A chain of heat-labile enterotoxin type IIB"
/codon_start=1
/transl_table=1
/product="enterotoxin"
/protein_id="AAAS3285.1"
/db_xref="GI:576585"
/translacion="MARVIFEFISLFLISPELVANDYFRADSRTPDEVRRSGGLIPRG
ODEAYERGTPTNINLYDHARGATGNTNRYNDGYSTTTTLOAHLLGONMLGYNERY
IYVVAAPNLEFDVNGVIGRSPYSENEFYALGIPLSQITGWRVSPGALIEGGMHRN
RDYRDLFRGISAPNEDGYRIAGPDPGPAWMEVPEPFAFNCLPNNKASSDTTCA
SLTNKLSQHLADLRKKTKRFTLMTLSTINNDGFFSNGKDEL"
113..172
/gene="LT-IIB"
173..901
/gene="LT-IIB"
/note="A chain of heat-labile enterotoxin type IIB"
/product="enterotoxin"
/note="B chain of heat-labile enterotoxin type IIB"
/codon_start=1
/transl_table=1
/product="enterotoxin"
/protein_id="AAAS3286.1"
/db_xref="GI:576586"
/translacion="MSFKITIKAFVIMAAIVSVOAHGASQFPKDCNRTTASIVEGV
ELTKYISDINNNTGMYVVSSTGGWIRISRAKDYPDNVTMRKIMAAVLSGRVAV
MCASPSSPNVITMILELAE"
963..1259
/gene="LT-IIB"
/note="B chain of heat-labile enterotoxin type IIB"
/product="enterotoxin"
BASE COUNT 369 a 195 c 285 g 413 t
ORIGIN
Alignment_scores:
Quality: 682.00 Length: 216
Ratio: 3.875 Gaps: 0
Percent Similarity: 81.481 Percent Identity: 57.407
alignment_block:
US-09-528-682-1 x ECOETOXHL ..
Align seg 1/1 to: ECOETOXHL from: 1 to: 1262
6 TyrArgAlaAspSerArgProProAspGluLeuLysArgSerGlyLeu 22
:::|||||
182 TTCAGGCGGTGATTCGAGAACGCGCTGATGAAGTCAGACGCTCAGCGGGGTT 231
:::|||||
22 uMetProAlaGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsn 39
:::|||||
232 GATTCGAAGAGGTCAGATGAGCTTATGAAAGCTGACACCCCTATTAACA 281
:::|||||
39 LeuAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArgTyr 55
|||||
282 TTAACTATATGACATGCAGAGGATCTGCACACAGGAACTAGATAT 331
|||||
56 AspAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisLeuAl 72
:::|||||
332 AATGATGATATGATATCAACACAGACTTTGAGACAGGCTCATTAT 381
:::|||||

```

```

72 aglyInserIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIle 89
|||||
382 AGGCAAAATATGCTTGCGGTATATGACGACTACTATATTTATGCTTG 431
|||||
89 IatHraIaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSer 105
|||||
432 CTCACACCAAAATTTGTTGATGATAATGCGCTTTAGGACAGTADAG 481
|||||
106 ProHisProTyrGluGlnGlyValSerAlaLeuGlyGlyLeuProTyrSe 122
|||||
482 CCGTATTCACAAAGTGAATGATATGCTGATGAGTGGTATTCGCCGTGC 531
|||||
122 rGlnIleTyrGlyTyrPheArgValAsnPhgIleValIleAspGluArg 139
|||||
532 ACAATTAATTTGCGTGTATAGACTATCTTTCGCTGCTATAGAGGGGGAA 581
|||||
139 euHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIle 155
|||||
582 TGCATGCAACACAGCATTTACAGAAAGATTTATTTAGAGGCTTATCTGC 631
|||||
156 AlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisSe 172
|||||
632 GCTCTAATGAGAGATGCTATCAATTCGCTGATTCGCCGACGATTTCC 681
|||||
172 nAlaTPrArgGluGluProTPrIleHisAlaProGlnGlyCysGlyA 189
|||||
682 TCGGTGGAGAAAGTTCCCTGGAGAAATTTGGCTTAATCTTGTCTAC 731
|||||
189 snSerSerArgThrIleThrGlyAspThrCysAsnGluGluThrGlnAsn 205
|||||
732 CAATATATAAGCTTCCAGGATACCTACCTGCTTAAACAATAATAA 781
|||||
206 LeuSerThrIleTyrLeuArgGlyTyrGlnSerLysValLysArgGln 221
|||||
782 CTATCACAGCATGATTACCTGATTTAAGAAATATATAAGAGAAAA 829
|||||
seq_name: gb_ba:ECOLIT1A
seq_documentation_block:
LOCUS ECOLIT1A 1200 bp DNA linear BCT 26-APR-1993
DEFINITION E.coli heat-labile enterotoxin type IIA (LT-IIa) A and B genes,
complete cds.
ACCESSION M17894
VERSION M17894.1 GI:146671
KEYWORDS enterotoxin; heat-labile enterotoxin.
SOURCE E.coli (strain 5A53) DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (bases 1 to 1200)
REFERENCE
1 Pickett,C.L., Weinstein,D.L. and Holmes,R.K.
Genetics of type IIA heat-labile enterotoxin of Escherichia coli:
operon fusions, nucleotide sequence, and hybridization studies
JOURNAL J. Bacteriol. 169, 5180-5187 (1987)
MEDLINE 88032841
FEATURES
source
location/qualifiers
1..1200
/organism="Escherichia coli"
/db_xref="taxon:562"
60..839
/note="heat labile enterotoxin type IIA A"
/transl_table=1
/protein_id="AAAS24093.1"
/db_xref="GI:146672"
/translacion="MKIKVLEFVIFSFVSANPEFRASRTPEIRRAGGLIPRGQO
EAYERGTPTNINLYDHARGATGNTNRYNDGYSTTTTLOAHLLGONMLGYNERYIY
VVAAPNLEFDVNGVIGRSPYSENEFYALGIPLSQITGWRVSPGALIEGGMHRN
RDYRDLFRGISAPNEDGYRIAGPDPGPAWMEVPEPFAFNCLPNNKASSDTTCA
SLTNKLSQHLADLRKKTKRFTLMTLSTINNDGFFSNGKDEL"
829..1200
/note="heat labile enterotoxin type IIA B"
CDS

```



```

/codon_start=1
/transl_table=1
/protein_id="AA24094.1"
/db_xref="GI:146673"
/translation="MSSKIIIGAFVIMTGLISGVYAGSPHEPNTICONTADIVAV
QKKYIDVNTNTGIYVSTGTGCVWYIPGGRDIPDNFISGEIKRTMAAILSDTKVN
LCATSSSPHIMWMEIDRES"
BASE COUNT      373 a      175 c      294 g      358 t
ORIGIN

```

```

alignment_scores:
  Quality: 679.50      Length: 238
  Ratio: 3.595      Gaps: 2
  Percent Similarity: 79.412      Percent Identity: 54.622

```

```

alignment block:
US-09-528-682-1 x ECOLITIA ..

```

```

Align seg 1/1 to: ECOLITIA from: 1 to: 1200

```

```

6  TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 172
123 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 172
22  TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 172
173 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 222
39  TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 55
223 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 272
56  TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 72
273 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 322
72  TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 89
323 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 372
89  TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 105
373 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 422
106 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 122
423 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 472
122 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 139
473 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 522
139 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 155
523 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 572
156 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 172
573 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 622
172 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 189
623 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 672
189 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 205
673 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 722
206 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 222
723 TTTTAAAGGAGTCTCCAGAACCTGATGTAATTAAGACGTGCGGAGGCT 772

```

```

222 e...PheSerAspTyrGlnSerGluValAspIle.....TyrAsnArgI 236
:      :      :      :      :      :      :      :      :
773 GCGGATACGTTTTCATGAGCGAAGATGATTTATGTTGTCATGAG 822
236 leArgAspGluLeu 240
|||||
823 AAGAGATGAGCTC 836

```

```

seq_name: gb_ba:ECOELT

```

```

seq_documentation_block:

```

```

LOCUS      ECOELT      774 bp      DNA      linear      BCT 25-JUN-1996
DEFINITION E.coli heat labile enterotoxin a 3' end and enterotoxin b (tox)
gene, complete cds.
ACCESSION J01646
VERSION J01646.1 GI:145828
KEYWORDS enterotoxin.
SOURCE      Escherichia coli (strain H10407) DNA.
ORGANISM    Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.

```

```

REFERENCE   1 (bases 190 to 292)
AUTHORS     Yamamoto,T., Tamura,T., Ryoji,M., Kaji,A., Yokota,T. and Takano,T.
TITLE       Sequence analysis of the heat-labile enterotoxin subunit B gene
            originating in human enterotoxigenic Escherichia coli
JOURNAL     J. Bacteriol. 152 (1), 506-509 (1982)
MEDLINE     83007048
REFERENCE   2 (bases 1 to 228)
AUTHORS     Yamamoto,T., Tamura,T., Yokota,T. and Takano,T.
TITLE       Overlapping genes in the heat-labile enterotoxin operon originating
            from Escherichia coli human strain
JOURNAL     Mol. Gen. Genet. 188 (2), 356-359 (1982)
MEDLINE     83114628
REFERENCE   3 (bases 190 to 774)
AUTHORS     Yamamoto,T. and Yokota,T.
TITLE       Sequence of heat-labile enterotoxin of Escherichia coli pathogenic
            for humans
JOURNAL     J. Bacteriol. 155 (2), 728-733 (1983)
MEDLINE     83265593
COMMENT     see also other loci beginning <ecoelt>.
FEATURES
  source

```

```

CDS

```

```

1..774
/organism="Escherichia coli"
/strain="H10407"
/species="Homo sapiens"
/db_xref="taxon:562"
<1..208
/codon_start=2
/product="enterotoxin A"
/protein_id="AA802981.1"
/db_xref="GI:1395121"
/translation="AMREEPWIHAPQGGCDSSTRTITGDCNFEQNLSTIYLRKQVS
KYRQIRSDYQSEVDIYNRIKNEI"
195..197
/gene="tox"
195..579
/gene="tox"
205..267
/gene="tox"
205..579
/gene="tox"

```

```

CDS

```

```

1..579
/gene="tox"
/codon_start=1
/transl_table=1
/product="enterotoxin B"
/protein_id="AA802982.1"
/db_xref="GI:1395122"
/translation="MNKVCYVLTALSSICAYGAPQSYTELCESEYRNQIYINDK
ILSYESMAKRENVITTFKSGATFQYEVGDSQHIDSQKAIIRMKDTLRITVLETK
IDKICVNNKTPNSIAISMEN"
268..576
/gene="tox"
/product="enterotoxin B"

```

```

mat_peptide

```

```

/gene="tox"
/product="enterotoxin B"

```


BASE COUNT 286 a 142 c 153 g 193 t
ORIGIN

alignment_scores:

Quality: 357.00 Length: 68
Ratio: 5.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 95.588

alignment_block:
US-09-528-682-1 x ECOELT ..

Align seg 1/1 to: ECOELT from: 1 to: 774

```

173 AAlATPARGGluGluProTPIleHISAlaProGlnGlyCysGlyAs 189
|||||
2 GCTTGAGAGAGAAACCTGGATTCATCATGCACACAGGTTGGAGA 51
|||||
189 nSerSerArgThrIleThrGlyAspThrCysAsnGluGluThrGlnAsnL 206
|||||
52 TTTCATCAAGAACATTCACAGCTGACTTGTAAATGAGAGACCCAGAAATC 101
|||||
206 eUserThrIleTyrLeuArgLysGluTyrGlnSerLysValLysArgGlnIle 222
|||||
102 TGACACACATATATCTCAGAAATATCATCAAAAGTTAAGAGCAGATA 151
|||||
223 PheSerAspTyrGlnSerGluValAspIleTyrAsnArgIleArgAspG 239
|||||
152 TTTTCAGACTATCATGACAGAGTTGACATATATTAACAGAAATTCGGAATGA 201
|||||
239 uLeu 240
|||||
202 ATTA 205

```

seq_name: gb_ba:ECLT87

seq_documentation_block:

LOCUS ECLT87 595 bp DNA linear BCT 28-OCT-1996
DEFINITION E.coli LT87 gene for heat-labile enterotoxin.
ACCESSION X83966
VERSION X83966.1 GI:1648865
KEYWORDS heat-labile enterotoxin; LT 87 gene.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE

1 (bases 1 to 595)
Germani, Y. and Desperrier, J.M.

Nucleotide sequence variations in genes encoding heat labile enterotoxins of Escherichia coli isolated in South Pacific

JOURNAL

2 (bases 1 to 595)
Unpublished

Germani, Y.
Direct Submission

Submitted (17-JAN-1995) Y. Germani, Institut Pasteur, 28 rue du Dr Roux, F-75724 Paris Cedex 15, FRANCE

Related sequence: M15361-3 (Yakamoto).
Location/Qualifiers

1. 595
/organism="Escherichia coli"
/strain="ETEC LT 87"
/db_xref="taxon:562"

205. 579
/gene="LT 87"

205. 579
/gene="LT 87"

205. 579
/gene="LT 87"

205. 579
/gene="LT 87"

205. 579
/gene="LT 87"

205. 579
/gene="LT 87"

205. 579
/gene="LT 87"

205. 579
/gene="LT 87"

BASE COUNT 226 a 105 c 113 g 151 t
ORIGIN
ILSYTESMAGKREMYLTPKSGATFOVEPGSHIDSOKKAIERMKDPLRLTYLTETK
IDKICVWNKPTNSTAISMEN"

alignment_scores:

Quality: 340.00 Length: 68
Ratio: 5.075 Gaps: 0
Percent Similarity: 98.529 Percent Identity: 97.059

alignment_block:
US-09-528-682-1 x ECLT87 ..

Align seg 1/1 to: ECLT87 from: 1 to: 595

```

173 AAlATPARGGluGluProTPIleHISAlaProGlnGlyCysGlyAs 189
|||||
3 GCTTGAGAGAGAAACCTGGATTCATCATGCACACAGGTTGTGAAA 52
|||||
189 nSerSerArgThrIleThrGlyAspThrCysAsnGluGluThrGlnAsnL 206
|||||
53 TTTCATCAAGAACATTCACAGCTGACTTGTAAATGAGAGACCCAGAAATC 102
|||||
206 eUserThrIleTyrLeuArgLysGluTyrGlnSerLysValLysArgGlnIle 222
|||||
103 TGACACACATATATCTCAGAAATATCATCAAAAGTTAAGAGCAGATA 151
|||||
223 PheSerAspTyrGlnSerGluValAspIleTyrAsnArgIleArgAspG 239
|||||
152 TTTTCAGACTATCATGACAGAGTTGACATATATTAACAGAAATTCGGAATGA 201
|||||
239 uLeu 240
|||||
202 ATTA 205

```

seq_name: gb_ba:ECOELTA2

seq_documentation_block:

LOCUS ECOELTA2 165 bp DNA linear BCT 26-APR-1993
DEFINITION E.coli heat labile enterotoxin subunit a gene; a1/a2 junction.
ACCESSION K00433
VERSION K00433.1 GI:145829
KEYWORDS enterotoxin.
SOURCE Escherichia coli DNA.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

REFERENCE

1 (bases 1 to 165)
Spicer, E.K., Kavanaugh, W.M., Dallas, W.S., Falkow, S.,

Konigsberg, W.H. and Schaefer, D.E.
Sequence homologies between a subunits of Escherichia coli and

Vibrio cholerae enterotoxins
Proc. Natl. Acad. Sci. U.S.A. 78, 50-54 (1981)

81223767
see other loci beginning <ecoeelt>.

Location/Qualifiers
1. 165
/organism="Escherichia coli"
/db_xref="taxon:562"

30 c 35 g 41 t
about 300 bp after seg 1.

BASE COUNT 59 a 30 c 35 g 41 t
ORIGIN

alignment_scores:
Quality: 247.00 Length: 53
Ratio: 4.843 Gaps: 1
Percent Similarity: 96.226 Percent Identity: 90.566

alignment_block:
US-09-528-682-1 x ECOELTA2 ..

Align seg 1/1 to: ECOELTA2 from: 1 to: 165


```

BASE COUNT      306 a      139 c      150 g      266 t
ORIGIN
VERKICVWNKNTPHIAIAISMANN"

alignment_scores:
  Quality:      182.00      Length:      55
  Ratio:        3.872      Gaps:      0
  Percent Similarity: 85.455      Percent Identity: 60.000

alignment_block:
  US-09-528-682-1 x VCCTXAB2      ..

Align seg 1/1 to: VCCTXAB2 from: 1 to: 861

186 GTCYCGASnserserargthrilethrGLYASPTRCysnsglucl 202
|||||
10 GGTTGGGGAATGCTCCAGATCATCGATGAGTAATCTGCATGAGAA 59
|||||
202 uThrGlnAaLeuSerThrIleTyLeuArgSgluTyGlnSerIysVal 219
|||||
60 AACCCAACTCAGTACGTAAATTCCTTCGACGAATACCAATCTAAAGTTA 109
|||||
219 ysaArgGlnIlePheSerAspTyGlnSerGluValAspIleTyGlna 235
|||||
110 AAAGACAATATTTTTCAGGCTATCATCTGATATGATACACATATAGA 159
|||||
236 ILeArgSPGLuLeu 240
|||||
160 ATTAAGATGAATTA 174

seq_name: gb_ba:VCCTXAB

seq_documentation_block:
LOCUS      VCCTXAB      938 bp      DNA      linear      BCT 10-FEB-1999
DEFINITION V.cholerae (0139) ctxA and ctxB genes, 938bp.
ACCESSION X76390
VERSION X76390.1 GI:433856
KEYWORDS cholera toxin; CT2A protein; CTB protein; ctxA gene; ctxB gene.
SOURCE Vibrio cholerae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
1 (bases 1 to 938)
Lebens,M.
Direct Submission
Submitted (24-NOV-1993) M. Lebens, Dept. of Medical Microbiol. &
Immunol., Goeteborgs univ., Guldhedsgatan, 10a, 413 46 Goeteborg,
SWEDEN
2 (bases 1 to 938)
Lebens,M. and Holmgren,J.
Structure and arrangement of the cholera toxin genes in Vibrio
cholerae O139
JOURNAL FEWS Microbiol. Lett. 117 (2), 197-202 (1994)
94237453
FEATURES
location/Qualifiers
1..938
/organism="Vibrio cholerae"
/strain="4260B"
/isolate="serotype 0139"
/db_xref="taxon:666"
37..177
/gene="ctxA"
37..177
/gene="ctxA"
/codon_start=1
/transl_table=1
/product="CT2A cholera toxin A2 subunit"
/protein_id="CAA53974.1"
/db_xref="GI:433858"
/db_xref="SPRPMBL:057372"
/translation="MSNTCDKTSGLGVFLDEQSKVKKQYSGYOSDIDTNRKID
EL"

prim_transcript 174..548

```


gene /gene="ctxb"
174..548
/gene="ctxb"
174..548
CDS /gene="ctxb"
/codon_start=1
/transl_table=1
/product="CTB cholera toxin B subunit"
/protein_id="CA53973.1"
/db_xref="GI:433857"
/db_xref="SWISS-PROT:P01556"
/translation="MIKLFQVFFVLLSAYAGHPONITDCAEYHNTQIYTLNDK
IFSYESLAGKREMAITFEKNGAIFQVEVPSQIHDSQKKAIERKKDLRIATYLEAK
VEKLCVMNNKTPHAIATISMAN"

mat_peptide

BASE COUNT 330 a 151 c 165 g 292 t
ORIGIN

alignment_scores:
Quality: 182.00 Length: 55
Ratio: 3.872 Gaps: 0
Percent Similarity: 85.455 Percent Identity: 60.000

alignment_block:

US-09-528-682-1 x VCCTXAB ..

Align seg 1/1 to: VCCTXAB from: 1 to: 938

186 GlyCysGlyAsnSerArgThrIleThrGlyAspThrCysAsnGluI 202
|||||
10 GGTGTGGATGCTCCAGATCATCGATGATACCTGCGATGAAA 59
202 urhGlnAsnLeuSerThrIleuArgGluTyrgInserLysVal 219
:|||||
60 ACCCAAGCTAGTGTAAATTCCTTGACGAATACCAATCTAAAGTTA 109
219 ysArgGlnIlePheSerAspTyrgInserGluValAspIleTyraArg 235
|||||
110 AAGACAAATTTTTCAGCTATCATGTATGATACATATATAGA 159
236 IleArgAspGluLeu 240
|||||
160 ATTAAGATGATTA 174

seq_name: gb_pat:A02701

seq_documentation_block:

LOCUS A02701 795 bp DNA linear PAT 17-MAR-1993
DEFINITION Synthetic DNA sequence for enterotoxoid production.
ACCESSION A02701
VERSION A02701.1 GI:344645

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 795)
ARTIFICIAL SEQUENCE.

AUTHORS
TITLE NEW PLASMIDS AND BACTERIAL STRAINS AND PRODUCTION PROCESS THEREO F
JOURNAL Patent: WO 8302456-A 1 21-JUL-1983;
FEATURES
source Location/Qualifiers
1..795
/organism="synthetic construct"
/db_xref="taxon:32630"

misc_feature

1..795

BASE COUNT 207 a 180 c 193 g 174 t
ORIGIN

alignment_scores:

Quality: 166.00 Length: 43
Ratio: 4.486 Gaps: 0
Percent Similarity: 86.047 Percent Identity: 79.070

alignment_block:

US-09-528-682-1 x A02701 ..

Align seg 1/1 to: A02701 from: 1 to: 795

188 GlyAsnSerArgThrIleThrGlyAspThrCysAsnGluI 204
|||||
664 GGAATTCATCAAGACATCAAGAGTGTACTGTATGAGGAGACCA 713
204 nasLeuSerThrIleuArgGluTyrgInserLysValysArg 221
|||||
714 GAATCTGAGCACAATATATCTCAGGGAATATCAATCAAAAGTTAAGGC 763
221 IuIlePheSerAspTyrgInserGluVal 230
|||||
764 AGATATTTTCAGACTATCAGCTCAAGSTN 792

seq_name: gb_pat:A02702

seq_documentation_block:

LOCUS A02702 795 bp DNA linear PAT 17-MAR-1993
DEFINITION Synthetic DNA sequence for enterotoxoid production (reverse
complement).

ACCESSION A02702

VERSION A02702.1 GI:344646

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.
ARTIFICIAL SEQUENCE.

REFERENCE 1 (bases 1 to 795)

AUTHORS

TITLE NEW PLASMIDS AND BACTERIAL STRAINS AND PRODUCTION PROCESS THEREO F

JOURNAL Patent: WO 8302456-A 2 21-JUL-1983;
FEATURES
source Location/Qualifiers
1..795
/organism="synthetic construct"
/db_xref="taxon:32630"

misc_feature

1..795

BASE COUNT 174 a 193 c 180 g 207 t
ORIGIN

alignment_scores:

Quality: 166.00 Length: 43
Ratio: 4.486 Gaps: 0
Percent Similarity: 86.047 Percent Identity: 79.070

alignment_block:

US-09-528-682-1 x A02702/rev ..

Align seg 1/1 to reverse of: A02702 from: 1 to: 795

188 GlyAsnSerArgThrIleThrGlyAspThrCysAsnGluI 204
|||||
132 GGAATTCATCAAGACATCAAGAGTGTACTGTATGAGGAGACCA 83
204 nasLeuSerThrIleuArgGluTyrgInserLysValysArg 221
|||||
82 GAATCTGAGCACAATATATCTCAGGGAATATCAATCAAAAGTTAAGGC 33
221 IuIlePheSerAspTyrgInserGluVal 230
|||||
32 AGATATTTTCAGACTATCAGCTCAAGSTN 4

seq_name: gb_pat:A06082

seq_documentation_block:

LOCUS A06082 791 bp DNA linear PAT 25-MAY-1993


```

alignment_block:
us-09-528-682-1 x A06083/rev ..
Align seg 1/1 to reverse of: A06083 from: 1 to: 791

188 GtYAsuSerArqThrlleThrGlyAsprHcYsaSngluGlunHrgl 204
||||| :|||: |||||||||||||||
128 GGAAATTCATCAAGACAAATCAGACGTCTACTGTGAATAGAGAGACCAC 79
||||| :|||: |||||||||||||||

204 nasnLeuSerThrlleTyrlLeuArgGlutryGlnserLySvallysArg 221
||||| :|||: |||||||||||||||
78 GAATCGACACACAATATATCTCAGGGAAATATCATCAAAGAATTAAAGAGC 29
||||| :|||: |||||||||||||||

221 InllePheSerAspyrGln 227
||||| :|||: |||||||||||||||
28 AGATATTTTTCAGACTATVCAG 9
||||| :|||: |||||||||||||||

seq_name: gb_ba:VIBTOX

seq_documentation_block:
LOCUS VIBTOX 512 bp DNA linear BCT 26-APR-1993
DEFINITION Vibrio cholerae toxA and toxB genes for cholera enterotoxin
subunits A2 (gamma) and B.
ACCESSION K01170 GI:155296
VERSION K01170
KEYWORDS enterotoxin.
SOURCE Vibrio cholerae (El Tor biotype strain 62746) DNA, clone pCV002.
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 512)
AUTHORS Lockman,H. and Kaper,J.B.
TITLE Nucleotide sequence analysis of the A2 and B subunits of Vibrio
cholerae enterotoxin
J. Biol. Chem. 258, 13722-13726 (1983)
MEDLINE 84061784
COMMENT The termination codon of subunit A2 overlaps the initiation codon
of subunit B by 4 bp. The correct translation of the 2 subunits is
the result of a shift in reading frames at this junction. Previous
protein structural analysis has shown that the A2 subunit is bound
to a precursor of the A subunit. Therefore, the 'met' at position 1
is not an initiator.
FEATURES
source location/Qualifiers
1..512
/organism="Vibrio cholerae"
/db_xref="taxon:566"
1..141
/note="enterotoxin subunit A2"
/codon_start=1
/transl_table=1
/protein_id="AAA27572.1"
/db_xref="GI:155297"
/translational="MSNTCDKETSLSGVFLDEYQSKVKRQIFSGYSIDITHNRKIKEL"
138..512
/note="enterotoxin subunit B"
/codon_start=1
/transl_table=1
/protein_id="AAA27573.1"
/db_xref="GI:155298"
/translational="MKIRGEVFPTVLSSAYAHGTPONTDLCAEYHNQIYTLDNK
IFSTESLACKREMAITTFKNGALFQVEVPDSOHIDSOKKAIEEMKDTLRATLYLTAKK
VERLCVVNNKTPIHAIAISMVN"
141..200
/note="enterotoxin subunit B signal peptide"
201..509
/product="enterotoxin subunit B mature peptide"
BASE COUNT 199 a 74 c 88 g 151 t
ORIGIN 188 bp upstream of NdeI site.

Alignment_scores:
Quality: 157.00 Length: 44
Ratio: 3.925 Gaps: 0

```


Percent Similarity: 90.909 Percent Identity: 65.909

Alignment block:
US-09-528-682-1 x VIBTOX ..

Align seg 1/1 to: VIBTOX from: 1 to: 512

```

197 AASPThCvASnglUgIuThrGlnAsnLeuSerThIleTyrlEuArgcl 213
   :::::::::::::::::::::::::::::::::::::::::::: ||
7  AATACTTGCGATGAAAAAACCAAGCTGAGGTGTAATAATCCCTTGACGA 56
   |||||||
213 uTyrgInserLysValIysArgGlnIlePheSerAspTyrgInserGluV 230
   |||||||
57 ATACCAATCTAAAGTTAAAGACAATATTTTCAGGCTATCATCTGATA 106
   |||||||
230 aAspIleTyrsnArGllleArGaspGluLeu 240
   ::|||
107 TTGATACACATATATAGATTAAAGATGAATTA 138

```

seq_name: gb_ba:VIBCTA2

seq_documentation_block:

```

LOCUS      VIBCTA2          251 bp    DNA          linear   BCT 26-APR-1993
DEFINITION Vibrio cholerae cholera toxin gene: subunit A gamma-chain 3' end 6
ACCESSION  J01847.1 GI:155157
VERSION    J01847
KEYWORDS   cholera toxin.
SEGMENT    2 of 2
SOURCE     Vibrio cholerae (strain El Tor 1621) DNA.
ORGANISM   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE  1 (bases 1 to 251)
AUTHORS   Gennaro,M.L. and Greenaway,P.J.
TITLE     Nucleotide sequences within the cholera toxin operon
JOURNAL   Nucleic Acids Res. 11, 3855-3861 (1983)
MEDLINE   83246519
FEATURES   Location/Qualifiers
            source          1..251
                        /organism="Vibrio cholerae"
                        /db_xref="taxon:66"
BASE COUNT      96 a      38 c      38 g      79 t
ORIGIN

```

Alignment_scores:

Quality: 150.00 Length: 44
Ratio: 3.846 Gaps: 0
Percent Similarity: 88.636 Percent Identity: 63.636

Alignment_block:
US-09-528-682-1 x VIBCTA2 ..

Align seg 1/1 to: VIBCTA2 from: 1 to: 251

```

197 AASPThCvASnglUgIuThrGlnAsnLeuSerThIleTyrlEuArgcl 213
   :::::::::::::::::::::::::::::::::::::::::::: ||
13 AATACTTGCGATGAAAAAACCAAGCTGAGGTGTAATAATCCCTTGACGA 62
   |||||||
213 uTyrgInserLysValIysArgGlnIlePheSerAspTyrgInserGluV 230
   |||||||
63 ATACCAATCTAAAGTTAAAGACAATATTTTCAGGCTATCATCTGATA 112
   |||||||
230 aAspIleTyrsnArGllleArGaspGluLeu 240
   ::|||
113 TTGATACACATATATAGATTAAAGATGAATTA 144

```

seq_name: gb_ba:ECOPHLEGS

seq_documentation_block:

```

LOCUS      ECOPHLEGS          900 bp    DNA          linear   BCT 26-APR-1993
DEFINITION Escherichia coli heat-labile enterotoxin gene, 5' cds.
ACCESSION  M61015

```

VERSION M61015.1 GI:147190
KEYWORDS heat-labile enterotoxin.
SOURCE E.coli, DNA.
ORGANISM Escherichia coli

REFERENCE 1 (bases 1 to 900)
AUTHORS Trachman,J.D. and Meas,W.K.
TITLE Characterization of the control region of the porcine heat-labile
JOURNAL enterotoxin in Escherichia coli
FEATURES Unpublished (1991)

```

JOURNAL    1 (bases 1 to 900)
TITLE      Characterization of the control region of the porcine heat-labile
SOURCE     enterotoxin in Escherichia coli
            Unpublished (1991)
            Location/Qualifiers
            source          1..900
                        /organism="Escherichia coli"
                        /strain="K-12"
                        /db_xref="taxon:562"
                        /product="heat-labile enterotoxin"
                        /protein_id="AA24335.1"
                        /db_xref="GI:147191"
                        /translation="MKNTFFIFLLASPLYANGDRLYRADSRPDEIKRSGL"
BASE COUNT      227 a      213 c      181 g      279 t
ORIGIN

```

CDS

```

CDS
mRNA
722..>900
781..>900
/codon_start=1
/transl_table=11
/product="heat-labile enterotoxin"
/protein_id="AA24335.1"
/db_xref="GI:147191"
/translation="MKNTFFIFLLASPLYANGDRLYRADSRPDEIKRSGL"
BASE COUNT      227 a      213 c      181 g      279 t
ORIGIN

```

Alignment_scores:

Quality: 117.00 Length: 22
Ratio: 5.318 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-528-682-1 x ECOPHLEGS ..

Align seg 1/1 to: ECOPHLEGS from: 1 to: 900

```

1  AAsngIyAspArgLeuTyraGalaAspSerArgProAspGluLeu 17
   |||||||
835 AATGCGACAGATTATACCGTCTGACTGACAGCCCAAGATGAATAA 884
   |||||||
17  sArgSerGlyGlyLeu 22
   |||||||
885 ACGTTCGCGAGGCTT 900

```

seq_name: gb_pat:A69697

seq_documentation_block:

```

LOCUS      A69697          2724 bp    DNA          linear   PAT 07-MAY-1999
DEFINITION Sequence 15 from Patent WO9807864.
ACCESSION  A69697
VERSION    A69697.1 GI:4774310
KEYWORDS   unidentified.
SOURCE     unidentified.
ORGANISM   unidentified.
            unclassified.
REFERENCE  1 (bases 1 to 2724)
AUTHORS   Foster,K.A., Quinn,C.P. and Shone,C.C.
TITLE     RECOMBINANT TOXIN FRAGMENTS
JOURNAL   Patent: WO 9807864-A 15 26-FEB-1998;
            FOSTER KEITH ALAN (GB)
FEATURES   Location/Qualifiers
            source          1..2724
                        /organism="unidentified"
                        /db_xref="taxon:32644"
                        /product="unnamed"
                        /product="unnamed"
            mat_peptide     1..2724
            CDS              1..2724
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CA842522.1"

```



```

      /db_xref="GI:4774311"
      /translation="MOPVNRQNPNNKPDVNGVDLAVIKIPNAGMOQPYPARKINKINMY
      IPEDDITFNEBGLDLPPEAKOVPAVSYYTSTYSTIDNEKNOKILGVTKEPERYSAD
      LGRRLTSLIVRGILEPMWGGSTIDELKAKIDINCINVIOPDGSYRSEELNLVIIGPSADI
      IOECSFGHEHVLNPNNGSGSTOYIFNSDPFTFFGESLEVDNPNLLGAGKATDPA
      VTLTAHELIIAAGHRLYCIAIDPNRYKATFNAAVEMSGLEVSFEELPTFGGHDAKFIDS
      LOEHREPLYYNKRKFOJASTLNKAKSVGTASALOYKANVEKEKYLISEDTSGKFSVD
      KIKDKXIKMLTEIYTEEDNVFKPKRNKTYLNEDEGAVERKINVPKVNTIYDGFMLE
      RNNLTANPENGOMETENNMATPKIKNTGCFEPKKILCVGCIITSRKMSDKGNKTE
      GRCDGALNDLCIKYNNMDLFPSPEDNFTDLNKGEETSIDTNIEMAEENISDDLIOQ
      YLTFENDEPENPESIENTLSNDIIIGOLEMPNIEREPNGKKEYELDKTMHYLRKOEF
      EHGSRIALNYSVMEALNPNSRVYTFFSSDYVKRKVNATEAMFLGWEDLVDFIDE
      TSEVSTCTADIADITIIPIYIGPALNIGNMKYKDFVVALIFSNAVILLEFIPETAIPA
      LGEPALSYIANKLVTOTLDNALSKNEKMDEVKYKIYTNMLKATVOJDILIRKKKK
      EAENOAEATRAKAIINOYNOYTEEREKNINENTIDDSSKLINESINKAMITVINKPLNOC
      SVSILNMSMTPYCGYKREDDEDSAKDKLLKTIYINRCITLIGOVDRILKDVYNNLTSDOI
      PFDLSKTVDMORLSTFTEYIKSRPOSVKVRKQIFPSGISODIDITHNRILKDEL"
BASE COUNT          974 a         418 c         509 g         823 t
ORIGIN

alignment_scores:
    Quality:   109.50           Length:       42
    Ratio:     3.318            Gaps:        1
    Percent Similarity: 78.571   Percent Identity: 54.762

alignment_block:
US-09-528-682-1 x A69697 ..

Align seg 1/1 to: A69697 from: 1 to: 2724

200 AsnGLuGlUthrClnAsnLeuSerThrIleTyrrLeuArgGluTr..._G1 215
|||||::: ||| |||||::: ||
2596 AATCAAGAATTATTTACTACATTTACTGCATATATAATTAAGTCAGACCCTCA 2645

215 nserLySValLySarGlnIlePheserSprGrInserGluValAspr 232
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2646 ACTTAAGTAAAGACAACAATATTTTCCAGCGCTATCATCTGATATTGATA 2695

232 leryrAsnarGllleaArspGluLeu 240
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2696 CACATATATGAATTAAGATGATTA 2721

seq_name: gb_pat:BD009884

seq_documentation_block:
LOCUS BD009884 2724 bp DNA linear PAT 31-JAN-20022
DEFINITION Recombinant toxin fragments.
ACCESSION BD009884
VERSION BD009884.1 GI:18638257
KEYWORDS JP 2001502890-A/8.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 2724)
AUTHORS Shone,C.C., Quinn,C.P. and Foster,K.A.
TITLE Recombinant toxin fragments.
JOURNAL Patent: JP 2001502890 A 8 06-MAR-2001;
MICROBIOLOGICAL RESEARCH AUTHORITY CAMR,THE SPEYWOOD LABORATORY LTD
COMMENT OS unidentified
PN JP 2001502890-A/8
PD 06-MAR-2001
PF 22-AUG-1997 JP 1998510524
PR 23-AUG-1996 GB 9617671.4 13-DEC-1996 GB 9625996.5 PT
CLIFFORD CHARLES SHONE,CONRAD PADURA QUINN,KEITH ALAN FOSTER PC
CIN15/31,CIN1/21,C1P21/02,C07K14/33,A61K39/08 CC
Strandedness: Double;
CC Topology: Linear;
FH key Location/Qualifiers
FT CDS 1..2724.
Location/Qualifiers
1..2724
/organism="unidentified"

```

```

BASE COUNT      974 a      418 c      509 g      823 t      /db_xref="taxon:32644"
ORIGIN

alignment_scores:
  Quality:      109.50      Length:      42
  Ratio:        3.318      Gaps:      1
  Percent Similarity: 76.571      Percent Identity: 54.762

alignment_block:
  US-09-528-682-1 x BD009884      ..

Align seg 1/1 to: BD009884 from: 1 to: 2724

200 AaAGlUGlUThrGlaSnLeuSerThrIleTyrlEuArGluTyr...GI 215
|||||.....:::|||||.....:
2596 AATCAAGATTTATTTACTGATTTACTGATTAATTAAGTCTAGGCCCTCA 2645

215 nserIysValIysArGlnIlePheSerAspTYrGlnSerGluValAspI 232
|||||.....:|||||.....:|||||.....:
2646 ATCTAAAGCTTAAAGCAATAATTTTCAGCGCATCATCATGTATATGATA 2695

232 lETyRAsnArGIlleArGAspGluLeu 240
:::|||||.....:|||||.....:
2696 CACATTAATGAATTTAAGATGAATTA 2721

seq_name: gb_ba:VIBCTA1

seq_documentation_block:
LOCUS      VIBCTA1      255 bp      DNA      linear      BCT 26-APR-1993
DEFINITION  Vibrio cholerae cholera gene: subunit A alpha-chain 5' end.
VERSION     J01846
ACCESSION   J01846.1 GI:155156
KEYWORDS    cholera toxin.
SEGMENT     1 of 2
SOURCE      Vibrio cholerae (strain El Tor 1621) DNA.
ORGANISM    Vibrio cholerae
REFERENCE   1 (bases 1 to 255)
AUTHORS     Gennaro,M.L. and Greenaway,P.J.
TITLE       Nucleotide sequences within the cholera toxin operon
JOURNAL     Nucleic Acids Res. 11, 3855-3861 (1983)
MEDLINE     83246519
COMMENT     clean copy provided by authors.
FEATURES
  source
    1..255
    Location/Qualifiers
      /organism="Vibrio cholerae"
      /db_xref="taxon:566"
BASE COUNT      85 a      23 c      38 g      109 t
ORIGIN

alignment_scores:
  Quality:      108.00      Length:      23
  Ratio:        4.909      Gaps:      0
  Percent Similarity: 95.652      Percent Identity: 86.957

alignment_block:
  US-09-528-682-1 x VIBCTA1      ..

Align seg 1/1 to: VIBCTA1 from: 1 to: 255

1 AaAGlYAspArGleuTyrrArGAlaAspSerArGProPrCAspGluIleTy 17
|||||.....:|||||.....:|||||.....:
187 AATGATGATTAAGTTATATCGGCGCATCTAGACCTCTGATGAATAATAA 236

17 sArGSeRGIyGlyLeuMet 23
|:::|||||.....:
237 GCAGTCAGGTGCTTATG 255

seq_name: gb_pat:103425

```


TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 167751)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (20-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Feb 13, 2002 this sequence version replaced g1:18581385. ----- Genome Center

Karlsson, E., Kelly, S., Khan, U., King, L., Korvath, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Lonsdale, H., Lozadó, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Mashwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Messay, E., Mawlin, E., McLeod, M. P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokoko, S., Ogun, M., Okunribido, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojchokan, I., Rolfe, M., Ruiz, S., Savary, G., Scheer, S., Scott, G., Shen, H., Shoohat, N., Sisson, I., Sultongren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tameris, A., Tameris, K., Tang, H., Tansley, J., Taylor, C., Taylor, T., Terfroid, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

Title: Direct Submission
 Reference: 2 (phases 1 to 167751)
 Authors: Worley, K.C.
 Journal: Direct Submission
 Submitted (20-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Feb 13, 2002 this sequence version replaced gi:18581385.

COMMENT

 Center: Baylor College of Medicine
 Web site: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: HDXK
 Center clone name: RP11-372M20

 Summary Statistics
 Sequencing vector: Plasmid; M77769
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 163427 bases at least Q40
 Consensus quality: 165469 bases at least Q30
 Consensus quality: 166917 bases at least Q20
 Estimated insert size: 167289; sum-of-coverage estimation
 Quality coverage: 0x in Q20 bases; agarose-gel estimation
 Quality coverage: 5x in Q20 bases; sum-of-coverage estimation

 NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 NOTE: This is a 'working draft' sequence. It currently
 consists of 8 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 42446: contig of 42446 bp in length
 * 42447 gap of unknown length
 * 42547 contig of 28869 bp in length
 * 71416 gap of unknown length
 * 71516 contig of 27640 bp in length
 * 99156 gap of unknown length
 * 99256 gap of unknown length
 * 124809 contig of 23534 bp in length
 * 124909 gap of unknown length
 * 124910 contig of 22144 bp in length
 * 147054 gap of unknown length
 * 147154 contig of 7109 bp in length
 * 147154 gap of unknown length
 * 154263 gap of unknown length
 * 154363 contig of 6795 bp in length
 * 161157 gap of unknown length

* 161258 167751: contig of 6494 bp in length.
Location/Qualifiers
Source 1.167751
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3q"
/clone="RP11-372M20"
BASE COUNT 54466 a 30864 c 31135 g 50585 t 701 others
ORIGIN

alignment_scores:
Quality: 101.00 Length: 149
Ratio: 1.262 Gaps: 7
Percent Similarity: 53.691 Percent Identity: 24.832

alignment_block:
US-09-528-682-1 x AC107423 ..

Align seg 1/1 to: AC107423 from: 1 to: 167751

100 ValLeuGIValTYrSerProHISProTYrGIuGIuValSerAlaLe 116
|||||::: ||| |||:::|
45208 GTGCTGGGTACACTGGATTCATCCATGCAGACAGATCAACT..... 45249
116 uGIyGIyLeProTYrSerGlnIleTYrGIyTYrTYrGValAspPheG 133
|||||:::|
45250AGAGTATCCTTCG 45262

133 lYValIleAspGIu.....ArgLeuHISArgAsnArgGIuTYrArgAsp 147
:::|
45263 CTACATACCAAAAGCAATCAAAATGTATTAAATCTTAATTAAAGAGT 45312

148 ArgTYrTYrArgAsnLeuAsnIleAlaProAlaGIuAspGIyTYrArgLe 164
|||||::: ||| |||:::|
45313 TCCTATTATTAACTACTACTAGTAAACATGGGAAACTCC..... 45354

164 uAlaGIyPheProAspHisGlnAlaTrpArgGIuGIuProTrpIleH 181
||| ||| ||| |||:::|
45355CCAGGACAT.....TGGAGGGGCAAAATTCCTTG 45385

181 lSHISAlaProGInGIyCySGIyAsnSerArgTrHleTrHGIyAsp 197
:::|
45386 ATCAATPCCCAAGCAGCCAGCAACCAAGCAAA.....ACTGGACAA 45429

198 ThrCysAsnGIuGIuThrGlnAsnLeuSerThrIleTYrLeuArgGIuTY 214
||| ::::::::::|
45430 ACAGGATCACTCACTTAATAAT.....CATCTCCACAAACA 45467

214 rGInSerLYsValLYsArgGln.....IlePheSerA 225
:::|
45468 AGGAACAAAGTAGAGACACACCAACCAAGGAAAGATATTACAA 45517

225 sPTyrgInSerGIuValAspIleTYrAsnArgIleArGAspGIuLeu 240
:::|
45518 ACTACTCATCTCAAGAGAAATTAATATAGACTATATTAAGAGCTC 45564

/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA104418 +	83.00	133.60	42.93	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAC86686 +	77.00	133.39	44.12	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH42258 +	83.00	128.19	85.89	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF12965 -	77.00	131.29	44.66	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAG57017 +	82.50	137.32	26.63	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABA057017 -	77.00	133.57	55.66	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB15399 +	82.50	127.98	98.23	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77110 -	77.00	126.99	100.24	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB6486 +	82.50	127.44	94.56	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77110 -	77.00	126.96	100.57	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB15398 -	82.50	115.21	453.65	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF86479 +	77.00	125.97	114.22	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB106028 +	82.50	79.44	1.1e+03	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAD24907 +	77.00	125.97	114.22	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH41226 +	82.50	108.41	4.4e+03	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC38988 +	77.00	116.59	380.33	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABA05185 +	82.00	133.78	16.37	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC38988 +	77.00	116.54	494.49	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB19899 +	82.00	133.78	16.37	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF86848 +	77.00	112.38	652.45	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB19868 +	82.00	124.78	133.11	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AB102234 +	77.00	108.67	1.0e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB101988 +	82.00	119.26	270.05	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAB1477 -	77.00	95.14	6.0e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB101990 +	82.00	119.25	270.43	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAB1477 -	77.00	70.07	1.5e+05	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAX22900 +	81.50	132.54	49.18	16	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF68534 +	77.00	68.81	1.7e+05	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD20414 -	81.50	112.54	639.49	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF21612 +	76.50	118.98	279.82	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAT17142 +	81.00	168.33	0.4988	45	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAF26999 +	76.50	116.60	379.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV33005 +	81.00	140.82	17.01	66	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO14668 +	76.50	116.20	389.94	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABA05189 +	81.00	138.91	21.72	79	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAV74830 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA158909 +	81.00	128.25	85.24	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA510458 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AA511474 +	80.50	132.00	52.70	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA611342 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAT41895 +	80.50	131.60	55.44	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS04122 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS01503 +	80.00	164.92	0.7928	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA823900 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS01504 +	80.00	164.73	0.7928	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF89102 +	76.50	109.74	915.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.AAO36375 +	80.00	164.53	0.8122	54	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF89102 +	76.50	103.84	2.0e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAC66187 +	80.00	139.14	26.46	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA82329 +	76.50	94.94	6.1e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB05189 +	80.00	137.37	26.46	76	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABA05184 +	76.00	129.56	72.05	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF66478 +	80.00	134.63	37.62	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV7941 -	76.00	127.19	97.69	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAD24906 +	80.00	134.63	37.62	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF691926 +	76.00	124.95	130.13	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH29871 +	80.00	128.70	80.51	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC74927 +	76.00	124.22	142.86	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH29874 +	80.00	128.13	86.61	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC74927 +	76.00	115.11	449.38	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.AAH29734 +	80.00	120.92	218.37	38	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA28209 +	76.00	115.29	494.88	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV31717 -	80.00	119.61	258.33	43	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA28215 +	76.00	113.22	585.71	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV31274 +	80.00	118.53	296.67	48	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB123316 -	76.00	112.34	655.98	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59666 -	80.00	109.14	988.58	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAN22019 +	76.00	98.85	3.7e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV73805 +	80.00	99.05	3.6e+03	32	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC36924 +	75.50	121.57	260.72	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV19941 +	80.00	84.18	2.4e+04	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA159247 +	75.50	121.04	214.98	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH99122 +	79.50	125.57	10.25	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB117185 +	75.50	117.68	330.86	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF21608 -	79.50	73.64	9.3e+04	34	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB123092 +	75.50	112.79	619.30	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV43064 +	79.00	164.47	0.8189	45	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB111122 +	75.50	103.36	2.1e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAZ27860 +	79.00	164.47	0.8189	45	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAB17820 -	75.00	90.24	1.1e+04	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ27494 +	79.00	164.47	0.8189	45	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AA511110 +	75.00	157.95	1.89	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ27496 +	79.00	132.88	47.07	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AA511119 +	75.00	157.95	1.89	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH35154 -	78.50	130.41	64.60	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AA51121 +	75.00	125.53	120.79	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF26410 +	78.50	124.90	131.07	19	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AA38230 -	75.00	125.53	120.79	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO12693 +	78.50	118.87	284.05	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX43737 +	75.00	117.61	379.14	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS79145 +	78.50	115.20	454.75	50	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC42350 +	75.00	116.61	379.14	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV52275 -	78.50	105.42	1.6e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS89784 +	75.00	116.08	405.83	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS51470 +	78.00	49.88	1.5e+06	29	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA05737 +	75.00	115.02	485.09	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD15273 +	78.00	133.43	43.89	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAO702337 +	75.00	108.52	505.57	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD15275 +	78.00	133.43	43.89	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.AAO70231 +	75.00	107.14	1.3e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABA05188 +	78.00	133.43	43.89	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB107004 +	75.00	106.95	1.3e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD03000 +	78.00	130.50	63.89	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS78436 +	75.00	106.95	1.3e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF13247 +	78.00	127.34	95.76	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS79370 +	75.00	106.95	1.3e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.AAO35306 +	78.00	123.79	150.97	19	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS71042 +	75.00	105.57	1.6e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.AAO76763 +	78.00	123.79	150.97	19	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF79882 -	74.50	102.61	2.3e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK69081 +	78.00	110.35	84.6.23	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AA557174 +	74.50	126.78	102.90	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB110614 +	78.00	109.96	889.90	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABA05185 +	74.50	126.78	102.90	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA103357 +	78.00	99.78	3.3e+03	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA568450 +	74.50	120.33	235.54	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS28908 -	78.00	99.78	3.3e+03	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.AAN70101 +	74.50	120.17	240.31	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK89083 +	78.00	136.46	29.76	52	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB112321 +	74.50	119.94	247.51	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT13644 +	77.50	129.61	71.06	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH66202 +	74.50	119.88	249.52	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF11251 +	77.50	130.59	63.72	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF71801 +	74.50	119.48	262.50	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB124019 +	77.50	117.92	330.58	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA151570 +	74.50	119.28	269.18	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AB124014 +	77.50	109.10	993.95	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF71535 +	74.50	119.07	276.57	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS69875 +	77.50	102.03	2.3e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AA562279 +	74.50	116.69	375.34	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAT13635 +	77.50	79.62	4.3e+04	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA527831 +	74.50	108.06	1.1e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH64966 -	77.50	69.77	1.5e+05	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK73305 -	74.50	108.06	1.1e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AA511122 +	77.00	161.81	1.15	40	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK73490 -	74.50	102.68	2.3e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AA511133 -	77.00	161.81	1.15	40	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AA57303 -	74.50	102.17	2.4e+03	1
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AA511136 -	77.00	161.81	1.15	40					

/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAT74269+	70.50	107.94	1.2e+03	22	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS4951+	70.00	104.26	1.8e+03	22
/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAT77025+	70.50	107.94	1.2e+03	22	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS4981+	70.00	104.02	1.9e+03	22
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAZ3925+	70.50	106.44	1.4e+03	26	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS7486+	70.00	103.80	2.0e+03	26
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAL1918+	70.50	106.43	1.4e+03	26	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAL1940+	70.00	103.62	2.0e+03	26
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAB62011+	70.50	106.36	1.4e+03	26	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAL05162+	70.00	103.51	2.0e+03	26
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAS5835+	70.50	106.32	1.4e+03	26	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAL13014+	70.00	103.29	2.1e+03	26
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAS5853+	70.50	105.76	1.5e+03	22	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS7377+	70.00	102.92	2.2e+03	22
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLB06777+	70.50	105.76	1.5e+03	22	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS9908+	70.00	102.86	2.2e+03	22
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAS48343+	70.50	105.12	1.7e+03	29	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABIL15034+	70.00	102.76	2.2e+03	29
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABLT27586-	70.50	103.51	2.0e+03	34	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABIL26669+	70.00	102.46	2.3e+03	34
/SID5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAAT94651+	70.50	103.51	2.0e+03	34	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABIL26695+	70.00	102.01	2.5e+03	34
/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAT96897+	70.50	103.51	2.0e+03	34	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74824+	70.00	102.01	2.5e+03	34
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAT819382-	70.50	103.47	2.0e+03	33	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74283+	70.00	102.01	2.5e+03	33
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAS22678-	70.50	102.89	2.2e+03	38	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74825+	70.00	102.01	2.5e+03	38
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC3193+	70.50	102.50	2.3e+03	38	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74833+	70.00	102.01	2.5e+03	38
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAC49498+	70.50	102.03	2.5e+03	40	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78940+	70.00	102.01	2.5e+03	40
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAS18274+	70.50	101.46	2.6e+03	4	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS81661+	70.00	102.01	2.5e+03	4
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAT80098+	70.50	100.67	2.9e+03	4	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS84092+	70.00	102.01	2.5e+03	4
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAS81541+	70.50	100.67	2.9e+03	4	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAT33697+	70.00	101.97	2.5e+03	4
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	100.34	3.1e+03	4	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL14101+	70.00	101.65	2.6e+03	4
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78719+	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT10547+	70.50	95.32	5.0e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78933+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAB61317+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78719+	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT10547+	70.50	95.32	5.0e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78933+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAB61317+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78719+	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT10547+	70.50	95.32	5.0e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78933+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAB61317+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78719+	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT10547+	70.50	95.32	5.0e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78933+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAB61317+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78719+	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT10547+	70.50	95.32	5.0e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78933+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAB61317+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78719+	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT10547+	70.50	95.32	5.0e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78933+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAB61317+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78719+	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT10547+	70.50	95.32	5.0e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78933+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAB61317+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78719+	70.00	101.54	2.6e+03	57
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:AAAT10547+	70.50	95.32	5.0e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS78933+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:ABBL19382-	70.50	95.55	5.7e+03	72	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAB61317+	70.00	101.54	2.6e+03	72
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF28427+	70.50	99.74	3.3e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:AAAG61317+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAF9074+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAG35650+	70.00	101.57	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:AAAS51775+	70.50	99.44	3.4e+03	51	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS94487+	70.00	101.56	2.6e+03	51
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF32423+	70.50	99.09	3.6e+03	5	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS74305+	70.00	101.54	2.6e+03	5
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAZ20606+	70.50	98.32	4.0e+03	57	/SID5/gcgdata/genese				

/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB64212	69.50	103.85	1.9e+03	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA89141	69.00	73.65	9.3e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI06517	69.50	101.75	2.6e+03	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAA1501	69.00	73.64	9.3e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.AAO27486	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA88142	69.00	73.63	9.3e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.AAO33879	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA89467	69.00	73.28	9.7e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAT92868	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA58471	69.00	71.61	1.2e+05	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAT92701	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI19214	69.00	70.88	1.3e+05	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAX91114	69.50	101.13	2.8e+03	36	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA54868	69.00	66.32	1.8e+05	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO12520	69.50	97.50	4.4e+03	51	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAF22288	69.00	66.97	2.1e+05	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.AAO25446	69.50	97.27	4.5e+03	52	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAF21611	69.00	53.35	1.1e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB16638	69.50	96.50	5.0e+03	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAA68533	69.00	53.35	1.1e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT75715	69.50	95.91	5.4e+03	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAA15489	69.00	44.41	2.4e+06	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABI33333	69.50	92.99	7.8e+03	80	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA57637	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAL03113	69.50	89.49	1.2e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA27059	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV62160	69.50	89.13	1.3e+04	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA05691	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT73194	69.50	88.83	1.3e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA31306	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS96629	69.50	88.79	1.3e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI15627	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV62133	69.50	88.29	1.4e+04	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI37203	68.50	120.23	238.57	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS95959	69.50	87.18	1.7e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX61794	68.50	118.04	315.90	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB120416	69.50	82.71	2.9e+04	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI17549	68.50	116.86	367.20	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAA81520	69.50	79.46	4.4e+04	30	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX61793	68.50	116.12	404.17	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAO54386	69.50	78.07	5.3e+04	34	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAB24215	68.50	114.42	502.15	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.AAO54386	69.50	74.91	7.9e+04	46	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABS56142	68.50	113.65	554.22	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAP30757	69.50	74.67	8.1e+04	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS56142	68.50	110.46	835.07	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAP30757	69.50	65.52	2.6e+05	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.ABA06734	68.50	110.07	877.08	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAD25116	69.50	62.67	3.6e+05	15	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS41457	68.50	110.07	877.08	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAD25116	69.50	54.31	9.6e+05	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS41457	68.50	110.07	877.08	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH68532	69.50	54.31	9.6e+05	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT83192	68.50	110.01	884.43	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH68532	69.50	46.35	8.37	40	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAT81801	68.50	109.72	917.71	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAA51130	69.00	146.35	8.37	40	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV41921	68.50	109.28	970.99	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAA51130	69.00	146.35	8.37	40	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB06561	68.50	109.17	985.13	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT71262	69.00	119.13	274.64	5	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAB80337	68.50	109.04	1.0e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAT71262	69.00	119.12	278.44	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS40837	68.50	108.75	1.0e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAF1357	69.00	118.12	312.59	62	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS40837	68.50	107.27	1.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH53663	69.00	116.52	383.58	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS49400	68.50	107.27	1.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH53663	69.00	115.49	437.95	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX22899	68.50	107.16	1.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAC33084	69.00	114.29	505.37	91	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI61186	68.50	107.16	1.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAT67889	69.00	113.10	555.37	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAI59401	68.50	106.62	1.4e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	69.00	112.03	682.91	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAO46543	68.50	105.30	1.6e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX7550	69.00	110.29	853.65	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAO46543	68.50	104.86	1.7e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAT67889	69.00	110.29	853.65	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT67889	68.50	103.54	2.0e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	69.00	109.38	930.85	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	68.50	103.54	2.0e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	69.00	109.38	930.85	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	68.50	103.32	2.1e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	69.00	109.38	930.85	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	68.50	103.14	2.1e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	69.00	108.77	1.0e+03	15	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	68.50	102.72	2.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAN3090	69.00	107.91	1.2e+03	17	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	68.50	102.47	2.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO10060	69.00	107.54	1.2e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT67889	68.50	102.47	2.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03129	69.00	107.11	1.3e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO14455	68.50	101.40	2.7e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03129	69.00	106.36	1.4e+03	19	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS78939	68.50	100.83	2.9e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS78939	69.00	104.65	1.8e+03	23	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS82654	68.50	100.83	2.9e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS78939	69.00	104.57	1.8e+03	23	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS82654	68.50	100.83	2.9e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV5376	69.00	104.34	1.8e+03	25	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAS49236	68.50	100.16	3.1e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS49236	69.00	103.85	2.0e+03	25	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS88046	68.50	99.81	3.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAB25999	69.00	103.85	2.0e+03	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS88046	68.50	99.81	3.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB25999	69.00	103.27	2.1e+03	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.AAN0681	68.50	99.67	3.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH03636	69.00	102.35	2.4e+03	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.AAO14478	68.50	99.35	3.5e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH03636	69.00	101.34	2.6e+03	2	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS68181	68.50	98.83	3.7e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAC86628	69.00	100.86	2.9e+03	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX0279	68.50	98.41	3.9e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS66400	69.00	100.25	3.1e+03	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03577	68.50	98.21	4.0e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03577	69.00	99.78	3.3e+03	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.ABI15693	68.50	98.05	4.1e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI12438	69.00	99.16	3.6e+03	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI15693	68.50	97.65	4.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH54322	69.00	98.82	3.7e+03	4	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH98351	68.50	97.65	4.3e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAC7490	69.00	97.69	4.3e+03	42	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA53019	68.50	97.37	4.5e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI2598	69.00	95.19	5.9e+03	58	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX19491	68.50	96.96	4.7e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV16307	69.00	95.19	5.9e+03	58	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.ABI9832	68.50	95.62	5.6e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT79751	69.00	94.25	6.7e+03	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.AAO50953	68.50	95.16	5.9e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAT79751	69.00	94.20	6.7e+03	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB08295	68.50	92.45	8.4e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000A.DAT.AAH57445	69.00	94.20	6.7e+03	6	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV20464	68.50	91.27	9.8e+03	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV16307	69.00	94.15	6.8e+03	65	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT10105	68.50	88.78	1.3e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59340	69.00	87.46	1.6e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB11788	68.50	87.90	1.5e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59340	69.00	85.15	2.1e+04	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS87216	68.50	86.12	1.9e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59356	69.00	82.56	3.0e+04	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI03576	68.50	76.90	6.1e+04	
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59356	69.00	77.21	5.9e+04	3	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABI07768	68.50			

/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH54688 - 67.50 97.48 4.4e+03 3
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAO85414 + 67.50 97.46 4.4e+03 35
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAO66246 + 67.50 97.46 4.4e+03 35
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAO03609 + 67.50 97.46 4.4e+03 3
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH8969 + 67.50 97.06 4.7e+03 3
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH8969 + 67.50 97.06 4.7e+03 3
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH05233 + 67.50 96.61 4.9e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH05233 + 67.50 96.14 5.2e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH17507 + 67.50 96.01 5.3e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABH15138 + 67.50 95.85 5.4e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAD24312 + 67.50 95.82 5.5e+03 4
 /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK52035 + 67.50 95.65 5.6e+03 4

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:AAH50205

seq_documentation_block:
 ID AAN50205 standard; DNA; 777 BP.

AC AAN50205;

DT 30-OCT-1991 (first entry)

DE Sequence of the pig scours heat labile toxin (LT) LTA gene.

DE Pig scours vaccine; toxin; diarrhoea; ss.

OS E.coli NCIB 11932.

Key Location/Qualifiers

FT CDS 1..777

FT /*tag= a

PN EPI45486-A.

PD 19-JUN-1985.

PE 12-DEC-1984; 84EP-0308620.

PR 12-DEC-1983; 83GB-0033131.

PA (GLAX) GLAXO GROUP LTD.

PI Hayes MV, Harford S, Ross GW;

DR WPI: 1985-148358/25.

DR P-PSDB: AAP50190.

PT New toxoid as inactivated form of toxin for use in vaccines - is

PT obd. from organism transformed by gene

PS Disclosure; Fig 1; 61pp; English.

CC AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of

CC the site directed mutant SDM1 (see AAN50206) is inactive. The

CC inventors claim a vaccine prep. active against pig scours which

CC contains an inactivated LTA component, together with additional K88

CC antigens opt. with whole cells comprising the antigens or contg. the

CC inactivated LTA.

SO Sequence 777 BP; 255 A; 136 C; 164 G; 222 T; 0 other;

alignment_scores:

Quality: 1304.00 Length: 240

Ratio: 5.433 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AAN50205 ..

Align seg 1/1 to: AAN50205 from: 1 to: 777

1 AaNgLyAsPaRgLeuTYrArGaLaAsPSeRaRqPrOPrOAsPGLuLeY 17
 ||||||||||||||||||||||||||||||||||||||||||||||||

55 AATGGCAGCATTTATACCGTGTGACTGTAGACCCCAAGATTAATAA 104
 17 sAaNgLyAsPaRgLeuTYrArGaLaAsPSeRaRqPrOPrOAsPGLuLeY 34
 105 ACCTTCGGAGGCTTATGCGGAGGAGCATATGAGTACTTCGATAGAG 154
 34 lYhRgImeTaNiLeaSnLeuTYrAsPhIsaLaRgLyhRgImR 50
 155 GAACCTAAATGAATTAATTCATTATGATCAGCGAGAGCAACCAAC 204
 51 GlyPheValArGTYrAsPaRgLyTYrValSerThrSerLeuSerLeuAr 67
 205 GCCTTGTCAGATATGATGACGGATATGTTCCACTTCCTTAGTTGAG 254
 67 gSeRaLaHIsLeuAlaGlyInSerIleuSerGlyTYrSerThrTYR 84
 255 AAGTGCTCACTTAGCAGACAGCTATATATACAGATATCCACTTACT 304
 84 yTlETyRValIleAlaThrAlaProAsnMetPheAsnValAsnSPAl 100
 305 ATATATATGTATTAGCAGACACCAATATGTTATATTTATATGATGA 354
 101 LeuGlyValTYrSerProHisProTYrGluGluValSerAlaLeuG1 117
 355 TTAGCGGTATACAGCCCTCACCATATGAACAGAGGTTTCGCTTAGG 404
 117 yGlyIleProTYrSerGlnIleTYrGlyTYrTYrArGyLaSnPheGly 134
 405 TGAATATCCATATTCATCATATATGATGATGATGATGATTAATTTGGTG 454
 134 aLlLeaSPAluArGLeuHIsArGAsnArGluTYrArGAsPaRgTYrTYR 150
 455 TGATTGATGAAGCATTTACATCTGTAACAGGATATAGAGACCGATTAC 504
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArGLeuAlaGly 167
 505 AGAAATCTGAATATAGCTCCGCGAGAGGATGCTTACAGATTAGCAGTTT 554
 167 ePrOPrAsPhIsGlnAlaThrAlaTYrGluGluProThrIleHIsAlaP 184
 555 CCCACCGATCCACCAAGCTTGAGAGAAAGAACCTTGATTCATCTGAC 604
 184 roGlnGlyCysGlyAsnSerSerArGThrIleThrGlyAspThrCysAsn 200
 605 CACAAGGTGTGGAATTCATCAACAACATCACAGCGTGAATCTGTAT 654
 201 GluGluThrGlnAsnLeuSerThrIleTYrLeuArGluTYrGlnSerly 217
 655 GAGGAGACCCAGAAATCTGAGCAATATATCTCAGGGATATTCATCAMA 704
 217 sValLySaRgInIlePheSerAsPtyrGlnSerGluValAsPlyLeTYrA 234
 705 AGTTAAGAGCGAGATATTTTCAGACTATCATCATGACAGAGGTGCATATATA 754
 234 sRAsPaRgLeuTYrArGaLaAsPSeRaRqPrOPrOAsPGLuLeY 240
 755 ACAGAAATTCGGGATGAATTA 774

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:AAH50205

seq_documentation_block:

ID AAN50206 standard; DNA; 777 BP.

AC AAN50206;

DT 30-OCT-1991 (first entry)

DE Sequence of the pig scours heat labile toxin (LT) LTA gene of the

DE site directed mutant SDM1.

DE Pig scours vaccine; toxin; diarrhoea; ss.

OS E.coli NCIB 11932.


```

XX Key Location/Qualifiers
FH CDS 1..777
FT /*tag= a
FT misc_difference 235..237
FT /*tag= b
FT /*note= "rcc in native sq"

XX EPI45486-A.
XX
XX 19-JUN-1985.
XX
XX 12-DEC-1984; 84EP-0308620.
XX
XX 12-DEC-1983; 83GB-0033131.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX PI Hayes MV, Harford S, Ross GW;
XX
XX WPI: 1985-148358/25.
XX P-PSDB: AAP50191.
XX
XX New toxoid as inactivated form of toxin for use in vaccines - is
XX obtd. from organism transformed by gene
XX
XX PS Example: Fig 2; 61pp: English.
XX
XX CC AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of
XX CC the site directed mutant SDM1 (see AAN50206) is inactive. The
XX CC inventors claim a vaccine prepn. active against pig scour which
XX CC contains an inactivated LTA component, together with additional K88
XX CC antigens opt. with whole cells comprising the antigens or contg. the
XX CC inactivated LTA.
XX
XX SQ Sequence 777 BP; 255 A; 135 C; 164 G; 223 T; 0 other;

alignment_scores:
Quality: 1298.00 Length: 240
Ratio: 5.431 Gaps: 0
Percent Similarity: 99.583 Percent Identity: 99.583

alignment_block:
US-09-528-682-1 x AAN50206 ..

Align seg 1/1 to: AAN50206 from: 1 to: 777

1 AanglYAspArgLeuTYrArgAlaAspSerArgProAspGluLeuY 17
|||||
55 AATGGCGACAGATTAACCGTGTGACTTACAGCCCCGATGAATAA 104
|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAnGluTYrPheAspArg 34
|||||
105 ACGTCCGAGAGCTTATGCCAGAGGCATATAGTACTTCGATAGAG 154
|||||
34 lYrThGlnMetAnlLeaAnLeuTYrAspHisAlaArgGlyHrGlnThr 50
|||||
155 GAACTCAAATGAATATTAATCTTATATGATCAGCGAGAGAACACAAACC 204
|||||
51 GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeuAr 67
|||||
205 GCCTTGTGAGATATGATACGAGATATGTTTCACTTCCTTAGTTGAG 254
|||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYrT 84
|||||
255 AAGTGCTCACTTAGCAGACAGCTATATTATACAGATATTCACACTTACT 304
|||||
84 YrIleTYrValIleAlaThrAlaProAsnMetPheAnValAsnAspVal 100
|||||
305 ATATATATATGTATAGCGACAGACAAATATATGTTTAAATGATGTA 354
|||||
101 LeuGlyValTYrSerProHisProTYrGluGlnGluValSerAlaLeuGl 117

```

```

|||||
355 TTAGGGGTATACAGCCCTCACCATATGAACAGAGAGCTTCGGCTTAGG 404
|||||
117 YGlyIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPheGlyY 134
|||||
405 TGAATATACATATTTTCAGATATATGATGATGATCTGTAAATTTTGCTG 454
|||||
134 AlIleAspGluArgLeuHisArgAsnArgGluTYrTrgAspArgTYrTYr 150
|||||
455 TGATTGATGACGATTTACATCGTACACGGGAAATATAGAACCCGATTTAC 504
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGlyPh 167
|||||
505 AGAAATCTGAATATAGCTCCGGCAGAGATGTTACAGATTACAGAGTTT 554
|||||
167 eProProAspHisGlnAlaTrpArgGluGluProTrpIleHisAlaP 184
|||||
555 CCCACCGGATCACACAGCTTGAGAGAGAACCCCTGATTCATCATGCAC 604
|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
605 CACAAAGTTGTGGAAATTCATCAAGAACATCACAGGTGATCTGTAAAT 654
|||||
201 GluGluTYrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSerly 217
|||||
655 GAGAGAGACCCAGAAATCTGACACAAATATATCTCAGGGAATATCAACAA 704
|||||
217 sValLYsArgGlnIlePheSerAspTYrGlnSerGluValAspIleTYrA 234
|||||
705 AGTTAAGAGGAGATATTTTTCAGACTATATCAGCAGAGTTGACATATATA 754
|||||
234 sNArgIleArgAspGluLeu 240
|||||
755 ACAGAAATTCGGGATGAATTA 774

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA2001A.DAT: AAS01505
seq_documentation_block:
ID AAS01505 standard; DNA; 1514 BP.
AC AAS01505;
XX
XX 29-AUG-2001 (first entry)
XX
XX DE E. coli heat-labile enterotoxin (LT) mutant LTS63Y DNA.
XX
XX KW Heat-labile enterotoxin; LT; LTS63Y; LTdel110/112; mutant;
XX KW detoxified and immunologically active protein; ADP-ribosylation; Gs;
XX KW endotoxin; diarrhoea; ds.
XX
XX OS Escherichia coli strain K88ac.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT CDS 164..1311
XX FT /*tag= a
XX FT /*product= "LTS63Y mutant protein"
XX FT /*transl_except= (pos:938..939,aa:met)
XX FT /*note= "This codon has an apparent 1 nucleotide
XX deletion which alters the reading frame"
XX
XX WO200119998-A1.
XX
XX PD 22-MAR-2001.
XX
XX PD 15-SEP-1999; 99WO-KR00555.
XX
XX PR 15-SEP-1999; 99WO-KR00555.
XX
XX PA (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
XX
XX PI Park EJ, Kim JS, Chang J, Yum J, Chung S;
XX

```


US-09-528-682-1 x AAAS1106 ..

Align seg 1/1 to: AAAS1106 from: 1 to: 782

```

1 ASNGLYSPARGLEUTYRARGALASPSEARARPROPSGLIILELY 17
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 AACGAGACACAGCTCTATAGCGCTAGCGCCCAAGATGAGACAA 109
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
17 SARSGEGLYGLYNEUMETPROARGLYHISASNGLYTYRPHESAPRG 34
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
110 GAGGTCTGGAGGTCTCATGCCAAGGACACAAATGACTTCTATAGCG 159
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
34 LYTHRGLMETASNIASNEUTYRASPHISALAARGLYTHRGLNTHR 50
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
160 GAACCCAAATGACATCACCTCATGACCATGCTAGCGGACCAACT 209
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
51 GLYPHEVALARGTYRASPSPGLYTYRVALSERTHRSEUSERLEUAR 67
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
210 GCGTTTGTAGGTWGTATGATGATGTGTCCACCTCTTACCTTGAG 259
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
67 GSERALHISLEUALAGLYNSERILEUSERGLYTYRSETHRTYR 84
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
260 GTCGTGCACACTGSCAGACATCATCTCTCGAGATCTCCACCTACT 309
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
84 YRILEYRVALILEALFTHRALAPROASNETHASNAVALASNPVAL 100
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
310 ACATCTATGTGATGTCTACAGACACAAACATGTTCATGTGAATGATGTG 359
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
101 LEUGLYVALTYRSEPROHISPROTYRGLUGLNUVALSERALALEUGL 117
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
360 TTGGGATGTGATGACCTCACCCATATGACACAGAGTGCTGTGGG 409
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
117 YGLYLEPROTYRSEGINILETYRGLYTYRARGVALASNPHEGLYV 134
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
410 TGAATCCCATCTACCCAAATCTATGATGTATAGGTTGACTTGGTG 459
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
134 ALLIEASPGIUAHGLEUHHISARGASARGGLUTYRARGASPARGYTYR 150
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
460 TGATGTGATGAGAGGCTCCATAGAGATAGGAGATAGGAGCAGGTACTAT 509
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
151 ARGASNLEUSNIIEALAPROALGLASPGLYTYRARGLEUALAGLYPH 167
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
510 AGGACCTCAACATAGCTCCAGCAGAGATGTATAGTTGGCAGGTTT 559
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
167 EPROPSPHISGLNALATRPARGGLUGLUPTROTPRIEHSIHISALAP 184
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
560 CCCACAGACACACACAGCTGAGGAGGAGGCCCTGATCCACCATGCAC 609
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
184 ROGINGLYCYSGLYASNSESRARGTHRILETHRGLYASPRHRCYASN 200
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
610 CACAAGGTTGTGAGACTCCTCAAGGACATCACAGGTGACACTTGCAAT 659
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
201 GLUGLUTHRGLASNLEUSERTHRILETYRLEUARGGLUTYRGINSERLY 217
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
660 GAGAGAGCCCAAAACCTTAGACCATTTACCTTAGGAGATGCCAATCAA 709
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
217 SVALLYSARGGLNLEPHESERAPRYRGINSERLVALASPILETYRA 234
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
710 GGTGAAGAGCAAAATCTTCTCAGACTACCAATCAGAGGTGACATCTACA 759
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
234 SNARGILEARGASPGIULEU 240
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||
760 ATAGATTAGGAATGAACCT 779
  |||||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAS1147

seq_documentation_block:

ID AAAS1147 standard; CDNA; 782 BP.

AC AAAS1147;

XX

26-SEP-2000 (first entry)

XX

DE Plant-optimized E. coli LT-A K63 mutant coding sequence.
 XX
 KW Heat-labile toxin: LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
 KW adjuvant; anti-bacterial; S63K; ss.
 XX
 OS Escherichia coli.
 OS Synthetic.

Key	Location/Qualifiers
FT CDS	3..782
FT	/*tag= a
FT	/product= heat-labile_toxin-A
FT	sig_peptide 3..59
FT	/*tag= b
FT	mat_peptide 60..779
FT	/*tag= c

WO20037609-A2.

29-JUN-2000.

22-DEC-1999; 99WO-US30747.

22-DEC-1998; 98US-0113507.

(BOYC-) BOYCE THOMPSON INST PLANT RES.

(MASO/) MASON H S.

(ARNT/) ARNTZEN C J.

Mason HS, Arntzen CJ;

WPI: 2000-442653/38.

P-PSDB: AAY96647.

New polynucleotides encoding LT-A or CT-A polypeptides for the

transformation of plant cells, useful in immunogenic compositions to

elicit immune responses in animals

Example 2; Page -; 103pp; English.

This cDNA encodes a mutant S63K Escherichia coli heat-labile toxin (LT) A subunit (LT-A). The codon at nucleotide positions 246-248 was changed from the wild-type TCC to AAG which causes a change of serine to lysine in the mature protein at residue 63. The sequence contains plant-preferred codons and eliminates sequence motifs associated with spurious mRNA processing. A single codon insertion (GAG encoding valine) was made to accommodate the creation of a NcoI restriction site around the initiator methionine codon. Novel polynucleotides encode a mutant LT-A polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit (CT-A) polypeptide, which have reduced enzyme activity as compared to the wild-type LT-A or CT-A polypeptide and where at least one of the codons is altered to a plant preferred codon. The polynucleotide further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B subunit (CT-B). The polynucleotides are useful for the transformation of plant cells for the production of transgenic plants to produce edible vaccines, especially oral vaccines in transgenic plants for the prophylactic or therapeutic treatment against E. coli or V. cholerae. The mutant polypeptides are also useful as adjuvants.

Note: This sequence does not appear in the specification. It was constructed from the wild type LT-A cDNA shown in AAAS1106 which is given in Figure 1 of the specification.

Sequence 782 BP; 227 A; 184 C; 192 G; 179 T; 0 other;

alignment_scores:

Quality	Ratio	Length
1283.00	5.346	240
Percent Similarity: 100.000	Percent Identity: 97.917	Gaps: 0

alignment_block:

US-09-528-682-1 x AAAS1147 ..

Align seg 1/1 to: AAA51147 from: 1 to: 782

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluIleu 17
  |||
60 AACGGAGACAAAGCTCTATAGGCTGACTAGGCCCCAGATGATGATCAA 109
  |||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
  |||
110 GAGGTGTGAGGCTCTATGCCAAGGGGACACATGAGTACTTGTATGGG 159
  |||
34 1YThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||
160 GAACCCAAATGAACATCAACCTCTATGACCATGCTAGGGGAACCAACT 209
  |||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  |||
210 GGTCTGTGAGGTATGATGATGATGATGATGATGATGATGATGATGAT 259
  |||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrThr 84
  |||
260 GTCTGTCTCACTTGGCAGACAAATCCATCCCTCAGATGATCTCCACT 309
  |||
84 yrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||
310 ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359
  |||
101 LeuGlyValTyrSerProHisProTyrGlyGlnGluValSerAlaLeu 117
  |||
360 TTGGGAGTATATAGCCCTCACCCTATGAGCAAGAGTGTCTCTTTGGG 409
  |||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPheArgValAsnPheGly 134
  |||
410 TGGAAATCCCATCTCCCAATCTATGATGATGATGATGATGATGATGAT 459
  |||
134 aIlleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
  |||
460 TGATGTATGAGAGGCTCCATAGGAATAGGAGATAGGGGACAGTACTAT 509
  |||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
  |||
510 AGGAACCTCAACATAGCTCCACGACGAGATGATGATGATGATGATGAT 559
  |||
167 eProProAspHisGlnAlaThrArgGlyGluIleProTyrIleHisHis 184
  |||
560 CCCACACAGACACCAAGCCTGGAGGAGAGAGCCCTGATCCACCAATGC 609
  |||
184 roGlnGlyCysGlyAsnSerSerArgTyrIleThrGlyAspThrCysAsn 200
  |||
610 CACAAAGGTGTGGAGACTCCTCAAGACCATCAGAGTGCACACTTGCAT 659
  |||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGlyTyrGlnSer 217
  |||
660 GAGGACACCCAAACCTTAGCACCATCTACCTTAGCAAGACCAATCCAA 709
  |||
217 sValIleAspGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
  |||
710 GGGAAAGAGGCAATCTCTCAACATCAACATCAAGAGGTGGACATCTACA 759
  |||
234 snArgIleAspGluLeu 240
  |||
760 ATAGGATTAGGAATGAATC 779
  |||

```

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA51545

seq_documentation_block:

ID AAA51545 standard; cDNA: 782 BP.

AC AAA51545;

DT 26-SEP-2000 (first entry)

DE Plant-optimized E. coli LT-A R72 mutant coding sequence.

```

KW Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
KM adjuvant; anti-bacterial; A72R; SS.
XX Escherichia coli.
OS Synthetic.
XX
FH Key location/Qualifiers
FT CDS 3..782
FT sig_peptide /tag= a
FT sig_peptide /product= heat-labile_toxin-A
FT mat_peptide /tag= b
FT /tag= c
PN WO200037609-A2.
XX
PD 29-JUN-2000.
XX
PF 22-DEC-1999; 99WO-US30747.
XX
PR 22-DEC-1998; 98US-0113507.
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX
PI Mason HS, Arntzen CJ;
PI
PI WPI: 2000-442653/38.
PI P-PSDB: AAY96649.
DR
DR
PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
PS Example 4; Page -: 103pp; English.
XX
CC This cDNA encodes a mutant A72R Escherichia coli heat-labile toxin (LT)
CC A subunit (LT-A). The nucleotides at position 273-274 were changed from
CC the wild-type GC to AG which causes a change of alanine to arginine in
CC the mature protein at residue 72. The sequence contains plant-preferred
CC codons and eliminates sequence motifs associated with spurious mRNA
CC processing. A single codon insertion (GTC encoding valine) was made to
CC accommodate the creation of a NcoI restriction site around the
CC initiator methionine codon. Novel polynucleotides encode a mutant LT-A
CC polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit
CC (CT-A) polypeptide, which have reduced enzyme activity as compared to
CC the wild-type LT-A or CT-A polypeptide and where at least one of the
CC codons is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible
CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae.
CC The mutant polypeptides are also useful as adjuvants.
CC Note: This sequence does not appear in the specification. It was
CC constructed from the wild type LT-A cDNA shown in AAA51106 which is
CC given in Figure 1 of the specification.
XX
SQ Sequence 782 BP; 226 A; 185 C; 191 G; 180 T; 0 other;

```

alignment_scores:

Quality: 1282.00 Length: 240
Ratio: 5.364 Gaps: 0
Percent Similarity: 99.583 Percent Identity: 97.917

alignment_block:

US-09-528-682-1 x AAA51545 ..

Align seg 1/1 to: AAA51545 from: 1 to: 782


```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
  |||
60 AACGGAGACAAAGCTATAGGCTGACTAGAGCCCAAGATGAGATCA 109
  |||
17 sargSerGlyLeuMetProArgHisAsnGluTyrPhasParg 34
  |||
110 GAGGTGGAGGTCTCATGGCAAGGGACAAATGACTACTTGTATGAGG 159
  |||
34 LyrThrIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||
160 GAACCAATGAACATCAACCTCTATGACCATGCTAGGGGAACCAACT 209
  |||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuLeu 67
  |||
210 GGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 259
  |||
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
  |||
260 GTCTGCTCATTTGAGAGGACAAATCCTCTCAAGGATCTCAACCTACT 309
  |||
84 YrlIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||
310 ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359
  |||
101 LeuGlyValTyrSerProHisProTyrGluGluGluValSerAlaLeu 117
  |||
360 TTGGGAGGTATAGCCCTCACCCTATAGCAAGAGGTGTCTCTTGGG 409
  |||
117 YGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
  |||
410 TGAATCCCTACTCCCAATCTATGATGATGATGATGATGATGATGATG 459
  |||
134 aIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
  |||
460 TGAATGATGAGAGGCTCCATAGGAATAGGATAGGAGGACAGTACTAT 509
  |||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
  |||
510 AGGAACCTCACAATAGCTCCAGCAGAGAGATGTTATAGTTGGCAGGTT 559
  |||
167 eProProAspHisGlnAlaThrPargGluGluProTyrIleHisIleAla 184
  |||
560 CCCACCAAGACCAACCAAGCTGGAGGAGGAGGCTGATCCACCATGAC 609
  |||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  |||
610 CACAAGGTTGTGAGACTCCTCAAGGACCATACAGGTGACACTTGCAAT 659
  |||
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
  |||
660 GAGGAGACCCAAACCTTAGCACCATCTACTTAGGAAGATCAACCAATCCA 709
  |||
217 sValIlyAspGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
  |||
710 GGTGAAGAGCAATCTCTCAACACTACCAATACAGAGGTGACATCTACA 759
  |||
234 snArgIleArgAspGluLeu 240
  |||
760 ATAGATTAGAATGAATC 779

```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51544

seq_documentation_block:

ID AAA51544 standard; cDNA; 782 BP.

AC AAA51544;

DT 26-SEP-2000 (first entry)

XX Plant-optimized E. coli LT-A G192 mutant coding sequence.

XX Heat-labile toxin: LT-A; LT-B; mutant; transgenic plant; vaccine; oral;

KW adjuvant; anti-bacterial; R192G; ss.

```

XX Escherichia coli.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT CDS 3..782
FT sig_peptide /tag= a
FT /product= heat-labile_toxin-A
FT mat_peptide /tag= b
FT /tag= c
FT
PN WO200037609-A2.
PD 29-JUN-2000.
PR 22-DEC-1999; 99WO-US30747.
PR 22-DEC-1998; 98US-0113507.
PR
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
PI
PI Mason HS, Arntzen CJ;
PI
PI WPI: 2000-442653/38.
PI P-PSDB; AAY96648.
PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells; useful in immunogenic compositions to
PT elicit immune responses in animals
PT
PS Example 3; Page -; 103pp; English.
XX
XX This cDNA encodes a mutant R192G Escherichia coli heat-labile toxin (LT)
XX A subunit (LT-A). The nucleotide at position 631 was changed from the
XX wild-type A to G which causes a change of arginine to glycine in the
XX mature protein at residue 192. The sequence contains plant-preferred
XX codons and eliminates sequence motifs associated with spurious mRNA
XX processing. A single codon insertion (CTG encoding valine) was made to
XX accommodate the creation of a NcoI restriction site around the
XX initiator methionine codon. Novel polynucleotides encode a mutant LT-A
XX polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit
XX (CT-A) polypeptide, which have reduced enzyme activity as compared to
XX the wild-type LT-A or CT-A polypeptide and where at least one of the
XX codons is altered to a plant preferred codon. The polynucleotide further
XX comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
XX subunit (CT-B). The polynucleotides are useful for the transformation of
XX plant cells for the production of transgenic plants to produce edible
XX vaccines, especially oral vaccines in transgenic plants for the
XX prophylactic or therapeutic treatment against E. coli or V. cholerae.
XX The mutant polypeptides are also useful as adjuvants.
XX Note: This sequence does not appear in the specification. It was
XX constructed from the wild type LT-A cDNA shown in AAA51106 which is
XX given in Figure 1 of the specification.
SQ Sequence 782 BP; 224 A; 166 C; 192 G; 180 T; 0 other;

```

alignment_scores:

Quality: 1280.00 Length: 240
Ratio: 5.356 Gaps: 0
Percent Similarity: 99.583 Percent Identity: 97.917

alignment_block:

US-09-528-682-1 x AAA51544 ..

Align seg 1/1 to: AAA51544 from: 1 to: 782

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
  |||

```



```

60 AACGGAGACAGCTCTATAGGCTGACTCTAGGCCCCAGATGATCA 109
17 sarserglyglyleumetproarglyhisnglutyrpheasparg 34
110 GAGGCTGGAGGCTCTATGCCAAGGGGACAAATGATGACTTTGATAGG 159
34 lythrcinmetasnileasnleuTYRASPHisAlaArglythrcinThr 50
160 GAACCCAAATGAAACATCAACCTCTATAGACCATCTGGGAAACCAACT 209
51 GlyPheValArgTYRASPaspGlyTYRValSerThrSerLeuSerLeuAr 67
210 GGCTTTGTAGGTATGATGATGATGATGATGATGATGATGATGATGATG 259
67 gserAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTYRSerThrTYT 84
260 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
84 yrlletyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
310 ACATCTATGATGATGCTACAGACCAACATGTTCAATGTGAATGATGTG 359
101 LeuGlyValTYRSerProHisProTYRGlutGlnGluValSerAlaLeuG 117
360 TTGGGAGGTATAGCCCTCACCATATGAGCAAGAGTGTCTGCTTGGG 409
117 ygllyleProTYRserGlnIleTYRGlYTYRtyrArgValAsnPhgIyV 134
410 TGGAAATCCCATCTACTCCCAATCTATGATGATGATGATGATGATGATG 459
134 alileaspGluArgLeuHisArgAsnArgGluTYRArgAspArgTYRtyr 150
460 TGAATGATATAGAGGCTCCATAGGAATAGGAGATATAGGACAGATCTAT 509
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYRArgLeuAlaGlyPh 167
510 AGGAACCTCAACATAGCTCCAGACAGAGATGATGATGATGATGATGATG 559
167 eProProAspHisGlnAlaTYRArgGluGluProTYRleHisHisAlaP 184
560 CCCACACAGACCAACCAAGCTGAGGAGAGAGCCCTGGATCCACCAATGC 609
184 roGlnGlycysGlyAsnSerSerArgTYRleThrGlyAspThrCysAsn 200
610 CACAAGGTTGTGAGACTCTCCAGGACCATCAGAGCTACACTTCAAT 659
201 GluGluThrGlnAsnLeuSerThrIleTYRLeuArgGluTYRcInSerly 217
660 GAGGAGACCCAAACCTTAGCACATCTACCTTAGGAGAGTACCAATCCAA 709
217 sValysATrgGlnIlePheSerAspTYRcInSerGluValAspIleTYR 234
710 GGTGAAGAGGCAAAATCTTCTCAGACTACCAATCAGAGGTGACATCTACA 759
234 snArgIleArgAspGluLeu 240
760 ATAGGATTAGGAATGACTC 779
seq_name: /SIDS5/gcdata/geneseq/genesepn-emb1/NA2001A.DAT: AAS01506
seq_documentation_block:
ID AAS01506 standard; DNA; 1508 BP.
XX
AC AAS01506;
XX
XX 29-AUG-2001 (first entry)
XX
DE E. coli heat-labile enterotoxin (LT) mutant LTdel110/112 DNA.
XX
KW Heat-labile enterotoxin; LT; LTs63Y; LTdel110/112; mutant;
KW detoxified and immunologically active protein; ADP-ribosylation; Gs;
XX endotoxin; diarrhoea; ds.
XX

```

```

OS Escherichia coli strain K88ac.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 164..1305
FT     /tag= a
FT     /product= "LTdel110/112 mutant protein"
FT     /transl_except= (pos:932..933,aa:Met)
FT     /note= "this codon has an apparent 1 nucleotide
FT     deletion which alters the reading frame"
XX
XX MO200119998-A1.
XX
XX 22-MAR-2001.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX (MOGAM-) MOGAM BIOTECHNOLOGY RES INST.
XX
XX Park EJ, Kim JS, Chang J, Yum J, Chung S;
XX
XX WPI: 2001-281524/29.
XX
XX P-PSDB: AAU00507.
XX
XX New detoxified mutants of Escherichia coli heat-labile enterotoxin
XX useful as vaccine for preventing and treating diarrhoea, and as adjuvant
XX for antibody production.
XX
XX
XX Claim 7; Page 45-46; 48pp; English.
XX
XX The present sequence encodes for Escherichia coli heat-labile
XX enterotoxin (LT) mutant LTdel110/112. LTs63Y (AAU00506) and LTdel110/112
XX are two novel detoxified and immunologically active proteins (LT
XX mutants) derived by site-directed mutagenesis of the A1 subunit of wild
XX type LT. The substitution of Ser to Tyr at position 63 in LTs63Y blocks
XX NAD-binding. Deletion of Glu residues at positions 110 and 112 in
XX LTdel110/112 eliminate the enzymatic activity of Gs. The A1 subunit of
XX wild type LT catalyses ADP-ribosylation of Gs, a GTP-binding protein that
XX regulates cAMP levels. The resulting increase in cAMP is the cause of
XX diarrhoea in humans and animals e.g. pigs. The mucosal immunogenicities
XX of mutant heat-labile enterotoxins LTs63Y and LTdel110/112 were tested.
XX Groups of mice were immunised with LTs63Y or LTdel110/112. The control
XX groups received phosphate buffered saline (PBS) alone. The serum and
XX faecal antibody titres to LT were determined. The results showed that
XX mice immunised with LTs63Y or LTdel110/112 contained high and
XX comparable level of anti-LT antibodies in sera and faecal extracts
XX compared with those immunised with wild-type LT. The LT mutants are
XX useful as a vaccine for preventing and treating diarrhoea and as an
XX adjuvant for antibody production.
XX
XX Sequence 1508 BP; 506 A; 277 C; 295 G; 430 T; 0 other:
SQ

```

```

alignment_scores:
Quality: 1280.00 Length: 240
Ratio: 5.378 Gaps: 1
Percent Similarity: 99.167 Percent Identity: 98.750

```

```

alignment_block:
US-09-528-682-1 x AAS01506 ..

```

```

Align seg 1/1 to: AAS01506 from: 1 to: 1508

```

```

1 AsnGlyAspArgLeuTYRArgAlaAspSerArgProProAspGluIleTy 17
|||||
218 AARGGCGACAGATTAACCGTGTGACTCTAGACCCCGAGATAAATAA 267
17 sarserglyglyleumetproarglyhisnglutyrpheasparg 34
|||||
268 ACCTCCGAGGCTCTATGCCAAGGGGACCAATGATGATGATGATGATGAG 317

```



```

34  lYthGlmetsnllleasnleutyrrasphsalsalarglythrglnthr 50
318 GAACCTCAATGAATATATATCTTTATGATCAGCGAGAGAACACAAACC 367
51  GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
368 GCGTTTGTCAGATATGATGACGAGATATGTTCCACTTCCTTAGTTGAG 417
67  gSerAlaHisLeuAlaGlyGlnSerTlleuSerGlyTyrSerThrTyrT 84
418 AAGTGTCTACTAGCAGACAGACTATATATATCAGATATTCCTACTACT 467
84  yTlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
468 ATATATATGTTATATAGCAGACCAAAATATGTTAATGTTATATGATGA 517
101 leuGlyValTyrSerProHisProTyrGlnGlnGlnValSerAlaLeuG 117
518 TTAGGCGTATACAGCCCTCACCCTAT...CAGGTTTCGCGTTAGG 561
117 yGlylleProTyrSerGlnIleTyrGlyTyrPtyrArgValAsnPhgGlyV 134
562 TGGATATACATATCTCAGATATATGAGATGATGCTGTTATTTTGGTG 611
134 allleAspGlnArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
612 TGATTGATGATACGATTTACATCGTAAACAGGAAATATAGACCGGATATAC 661
151 ArgAsnLeuAsnIleAlaThrAlaProAlaGlnAspGlyTyrArgLeuAlaGlyph 167
662 ACAAATCTGAATATAGCTCCGCGAGAGATGTTACAGATTTAGCAGGTTT 711
167 eProProAspHisGlnAlaTyrPArgGlnGlnLeuProTyrPheHisAlaP 184
712 CCCACCGGATTCACCAAGCTTGAGAGAAACCCGATTCATCATGCCAC 761
184 roGlnGlyCysGlnAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
762 CACAAGGTTGTGAAATTCATCAAGAACATCAACAGGTCATCTGTAAT 811
201 GlnGlnThrGlnAsnLeuSerThrIleTyrLeuArgGlnTyrGlnSerly 217
812 GAGGAGACCCAGAAATCTGAGCAAAATATCTCAGGGAATATCAATCAAA 861
217 sValIysArgGlnIlePheSerAspTyrGlnSerGlnValAspIleTyrA 234
862 AGTTAAGAGCGAGATATTTTCAGACATCATCAGTCAGAGGTTGACATATATA 911
234 snArgIleArgAspGlnLeu 240
912 ACAGAAATTCGGGATGAAATTA 931
seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.AA042768
seq_documentation_block:
ID AA042768 standard; DNA, 711 BP.
XX
AC AA042768;
XX
DT 08-DEC-1993 (first entry)
XX
DE E.coli heat labile toxin subunit A coding sequence.
XX
KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KW ADP-ribosyltransferase activity; ss.
XX
OS Escherichia coli.
XX
FH Key location/Qualifiers
FT mat_peptide 1..708
FT /*tag= a
FT /note= "LT-A"

```

```

XX
PN WO9313202-A.
XX
PD 08-JUL-1993.
XX
PF 30-DEC-1992; 92WO-EP03016.
XX
PR 31-DEC-1991; 91IT-0M13513.
XX
PA (BIOC-) BIOCINE SCLAVO SPA.
XX
PI Domenighini M, Hol W, Pizsa M, Rappuoli R;
XX
DR WPI: 1993-227320/28.
XX
DR P-PSDB; AAR8728.
XX
PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli
XX
PS Disclosure; Fig 2; 60pp; English.
XX
CC This sequence encodes the A subunit of the heat labile toxin (LT-A)
CC of a strain of E.coli known to affect humans. The sequence was
CC published by Yamamoto et al, J.Biol. Chem., 259, 5037-5044.
CC Mutations at selected positions within this sequence have been found
CC to reduce toxicity (see AA051314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli.
XX
SQ Sequence 711 BP; 234 A; 123 C; 156 G; 198 T; 0 other:
XX
alignment_scores:
Quality: 1183.00 Length: 240
Ratio: 5.189 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 93.333
alignment_block:
US-09-528-682-1 x AA042768 ..
Align seg 1/1 to: AA042768 from: 1 to: 711
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGlnIlely 17
|||||
1 AATGGCAGACAGATTAATACCGTCTGACTGATAGACCCAGATGAATATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlnTyrPheAspArgG 34
|||||
51 ACGTTTCGCGAGCTTATGCCAGAGGT...AATGAGTACTTCGATAGAG 97
34 lYthGlmetsnllleasnleutyrrasphsalsalarglythrglnthr 50
|||||
98 GAACCTCAATGAATATATATCTTTATGATCAGCGAGAGAGAACACAAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||
148 GCGTTTGTCAGATATGATGACGAGATATGTTCCACTTCCTTAGTTGAG 197
67 gSerAlaHisLeuAlaGlyGlnSerTlleuSerGlyTyrSerThrTyrT 84
|||||
198 AAGTGTCTACTAGCAGACAGACTATATATATCAGATATTCCTACTACTA 247
84 yTlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
248 TATATATCTGTTATAGCA.....ATATGTTTAATGTTAATGATGA 288
101 leuGlyValTyrSerProHisProTyrGlnGlnGlnValSerAlaLeuG 117
|||||
289 ATTAGCGTATACAGCCCTCACCCTATATAGACGAGGTTTCGCGTTAGG 338
117 yGlylleProTyrSerGlnIleTyrGlyTyrPtyrArgValAsnPhgGlyV 134
|||||

```


CC E. Coli neat labile toxin, an example of a bacterial ADP-ribosylating
CC toxin. A mutant detoxified form of this protein is used in a parentera
CC adjuvant composition, which comprises the detoxified protein, at least
CC one selected antigen and optionally a pharmaceutically acceptable
CC (optionally topical) vehicle. The adjuvant composition can be
CC administered parenterally in conjunction with at least one antigen in
CC methods to immunise vertebrate subjects. The adjuvant has the ability
CC to enhance the humoral and cell-mediated immune responses elicited by
CC the antigen (e.g. by making the antigen more strongly immunogenic or
CC necessitating fewer/lower antigen doses). It can be administered
CC prior/subsequent to the antigen, and is preferably administered within
CC a short space of time to the same site: it can also be administered in
CC isolation from antigens as a boost following systemic or mucosal antigen
CC administration. Most preferably, the adjuvant is co-administered with
CC the antigen in the compositions and a pharmaceutically acceptable
CC carrier. The antigen may be derived from viruses, bacteria, parasites
CC and fungi or may be tumour antigens, self-antigens and allergens. The
CC compositions are therefore useful in the treatment and prevention of
CC e.g. viral diseases, allergic manifestations, diseases caused by
CC pathogens (e.g. bacteria or parasites), AIDS, autoimmune diseases
CC (e.g. Systemic Lupus Erythematosus), Alzheimer's disease and cancers.
CC The adjuvant can also be used to prepare antbodies against selected
CC antigen(s), useful e.g. for diagnostic purposes or for antigen
CC purification.

xx
xx
Sequence 711 BP: 234 A; 123 C; 156 G; 198 T; 0 other:

Alignment_scores:
Quality: 1183.00 Length: 240
Ratio: 5.189 Gaps: 2

alignment_block:

Align seg 1/1 to: AAV81595 from: 1 to: 711

1 234567891011121314151617181920212223242526272829303132333435363738394041424344454647484950515253545556575859606162636465666768697071727374757677787980818283848586878889909192939495969798991001011021031041051061071081091101111121131141151161171181191201211221231241251261271281291301311321331341351361371381391401411421431441451461471481491501511521531541551561571581591601611621631641651661671681691701711721731741751761771781791801811821831841851861871881891901911921931941951961971981992002012022032042052062072082092102112122132142152162172182192202212222232242252262272282292302312322332342352362372382392402412422432442452462472482492502512522532542552562572582592602612622632642652662672682692702712722732742752762772782792802812822832842852862872882892902912922932942952962972982993003013023033043053063073083093103113123133143153163173183193203213223233243253263273283293303313323333343353363373383393403413423433443453463473483493503513523533543553563573583593603613623633643653663673683693703713723733743753763773783793803813823833843853863873883893903913923933943953963973983994004014024034044054064074084094104114124134144154164174184194204214224234244254264274284294304314324334344354364374384394404414424434444454464474484494504514524534544554564574584594604614624634644654664674684694704714724734744754764774784794804814824834844854864874884894904914924934944954964974984995005015025035045055065075085095105115125135145155165175185195205215225235245255265275285295305315325335345355365375385395405415425435445455465475485495505515525535545555565575585595605615625635645655665675685695705715725735745755765775785795805815825835845855865875885895905915925935945955965975985996006016026036046056066076086096106116126136146156166176186196206216226236246256266276286296306316326336346356366376386396406416426436446456466476486496506516526536546556566576586596606616626636646656666676686696706716726736746756766776786796806816826836846856866876886896906916926936946956966976986997007017027037047057067077087097107117127137147157167177187197207217227237247257267277287297307317327337347357367377387397407417427437447457467477487497507517527537547557567577587597607617627637647657667677687697707717727737747757767777787797807817827837847857867877887897907917927937947957967977987998008018028038048058068078088098108118128138148158168178188198208218228238248258268278288298308318328338348358368378388398408418428438448458468478488498508518528538548558568578588598608618628638648658668678688698708718728738748758768778788798808818828838848858868878888898908918928938948958968978988999009019029039049059069079089099109119129139149159169179189199209219229239249259269279289299309319329339349359369379389399409419429439449459469479489499509519529539549559569579589599609619629639649659669679689699709719729739749759769779789799809819829839849859869879889899909919929939949959969979989991000100110021003100410051006100710081009101010111012101310141015101610171018101910201021102210231024102510261027102810291030103110321033103410351036103710381039104010411042104310441045104610471048104910501051105210531054105510561057105810591060106110621063106410651066106710681069107010711072107310741075107610771078107910801081108210831084108510861087108810891090109110921093109410951096109710981099110011011102110311041105110611071108110911101111111211131114111511161117111811191120112111221123112411251126112711281129113011311132113311341135113611371138113911401141114211431144114511461147114811491150115111521153115411551156115711581159116011611162116311641165116611671168116911701171117211731174117511761177117811791180118111821183118411851186118711881189119011911192119311941195119611971198119912001201120212031204120512061207120812091210121112121213121412151216121712181219122012211222122312241225122612271228122912301231123212331234123512361237123812391240124112421243124412451246124712481249125012511252125312541255125612571258125912601261126212631264126512661267126812691270127112721273127412751276127712781279128012811282128312841285128612871288128912901291129212931294129512961297129812991300

www.tuwatw.com

```

1 / satybceioylyucemecfiohiybiyhnsksnoiaiyifnchsbkayub 34
||| :::||| |||||

```

[illegible]

34 1y1n01m6c4s011e4s0u0e0y1n3p4n1s1a0a1g0y1m0v0m0m 20

30 GAAACACAAAGAAATATATATCTTATATGATCTGCGAGAGGATACACACAAATC 14

01 019fnevalnig y i nsphns pol y i val sei n n l sei deu sei deu ki v

|||||

[illegible][illegible]

23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 105

[illegible][illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhe 167
167 ePProAspHisGlnAlaIleTyrArgGluIlePheTyrPheHisAlaPhe 184
489 AGAATCTGAATATAGCTCCGCGAGAGAGATTACAGATTACAGAGTTT 488
489 CCCACCGGATCCACCAAGCTTGAGAGAGAAACCTGGATTCAATCATGAC 538
184 roGInGlyCysGlyAsnSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAGGTTTGAGATTCTCATCAAGAACATTCACAGGTGATTGTAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluIleTyrGlnSer 217
589 GAGGAGACCAGATCTGACACACATATATCTCAGGGAATATCAATCAAA 638
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
639 AGTTAAAGCGCAGATATTTTCAGACTATCAGTCAGAGTTGACATATATA 688
234 sNArgIleArgAspGluLeu 240
689 ACAGAAATTCGGAGATGAAATTA 708
seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AA051317

seq_documentation_block:
ID AA051317 standard: DNA; 711 bp.

XX AA051317;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-63 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KM ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

XX Key Location/Qualifiers
FT misc_difference 184..186
FT /tag= a
FT /note= "Wild-type TCT(Ser) mutated to AAG(Lys)"

PN WO9313202-A.

XX 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-OMI3513.

XX (BIOC-) BIOCINE SCLAVO SPA.

PI Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX MPI: 1993-227320/28.

DR P-PSDB: AAR44016.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al. J.Biol. Chem. 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-051326). The invention relates to

CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051317 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 bp; 236 A; 122 C; 157 G; 196 T; 0 other;

alignment_scores:
Quality: 1179.00 Length: 240
Ratio: 5.171 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 92.917

alignment_block:
US-09-528-682-1 x AA051317 ..

Align seg 1/1 to: AA051317 from: 1 to: 711

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLys 17
1 AATGGCGACAGATTATACCGTGCTGACTCTAGACCCCGAGATGAATATA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 ACCTTCCCGAGCTTATGCCAGAGT...AATGAGTCTTCGATGAGAG 97
34 LThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
98 GAACCTAAATGAATATTAATCTTTATGATCAGCCGAGAGAACACAAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
148 GCGTTGTCAGATATGATGACGAGATATGTTCCACTAAGCTTAGTTGAG 197
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
198 AAGTGTCTACTTGACGAGACAGATATATATATCAGATATTCCTACTTA 247
84 yrlLetyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATAGCA.....AATATGTTAATGTTAATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGlnGluValSerAlaLeuGln 117
289 AATAGCGTATACAGCCCTCACCATATGAACAGAGAGTTTCGCTTAGG 338
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
339 TGGAAATACCATATTCAGATATATGATGATGATCGTTAATTTGGTG 388
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
389 TGAATTATGACCATTTACATCGTAACAGGAAATATAGAACCGGTATTAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhe 167
439 AGAATCTGAATATAGCTCCGCGAGAGAGATTGTTACAGATTACAGAGTTT 488
167 ePProAspHisGlnAlaIleTyrArgGluIlePheTyrPheHisAlaPhe 184
489 CCCACCGGATCCACCAAGCTTGAGAGAGAAACCTGGATTCAATCATGAC 538
184 roGInGlyCysGlyAsnSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAGGTTTGAGATTCTCATCAAGAACATTCACAGGTGATTGTAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluIleTyrGlnSer 217
589 GAGGAGACCAGAAATCTGAGCAATATATCTCAGGGAATATCAATCAAA 638
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234

639 AGTTAAGGACGACATATTTTCACACTATCAGTCAGAGGTTGACATATATA 688

234 snArgIIleArgAspGluLeu 240

|||||

689 ACAGAAATTCGGGATGAATTA 708

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51325

seq_documentation_block:

ID AAQ51325 standard; DNA: 711 BP.

AC AAQ51325;

DT 08-DEC-1993 (first entry)

DE Encodes Glu-114 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

XX Key Location/Qualifiers

FT misc_difference 328..330 /tag= A

FT /note= "wild-type TCT(Ser) mutated to GAA(Glu)"

XX WO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992: 92WO-EP03016.

XX 31-DEC-1991: 91IT-0M13513.

XX (BIOC-) BIOFINE SCLAVO SPA.

XX Domenighini M, Hol W, Piazza M, Rappuoli R;

XX WI: 1993-227320/28.

XX P-PSDB; AAR44024.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX - useful as vaccines against infection by Vibrio cholerae and

XX enterotoxin producing Escherichia coli

XX Claim 3: Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile

XX toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto

XX et al, J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to

XX site-directed mutagenesis. Certain mutations were found to reduce

XX toxicity (see AAQ51314-051326). The invention relates to

XX immunogenic, detoxified LT-A proteins and their use in vaccines to

XX protect against enterotoxigenic E.coli. Sequence AAQ51325 is a

XX combination of the wild-type coding sequence and the mutagenic

XX primer sequence used to introduce the preferred mutation. (Amino

XX acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 236 A; 122 C; 157 G; 196 T; 0 other;

Alignment_scores:

Quality: 1179.00

Ratio: 5.171

Percent Similarity: 95.000

Percent Identity: 92.917

Alignment_block:

US-09-528-682-1 x AAQ51325

Length: 240

Gaps: 2

Percent Identity: 92.917

Align seg 1/1 to: AAQ51325 from: 1 to: 711

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17

|||||

1 AATGGGACGATATATACCGTCTGACTAGACCCACATGAAATATA 50

|||||

17 sarGserGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34

|||||

51 ACCTTTCCGGAGCTTATGCCAGAGGT...AATGACTACTGATAGAG 97

|||||

34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50

|||||

98 GAACTCAAAATGAATATTAATCTTATGATCATCCGAGAGGACCAAAAC 147

|||||

51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67

|||||

148 GGCTTGTCTCAGATATGATGACGGATATGTTCCACTCTCTTGTTGAG 197

|||||

67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84

|||||

198 AATGCGCACTTAGCAGACAGATATATATATCAGGATATCACTTACTA 247

|||||

84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100

|||||

248 TATATATCGTTATATAGCA.....AATATGTTTATATGTTATATGATGTA 288

|||||

101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117

|||||

289 AATAGCGTATACAGCCCTCACCATATGACAGAGGTTGACAGGTTAGG 338

|||||

117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134

|||||

339 TGGATATACCATATTCAGATATATATGATGATCGTTAATTTTGGTG 388

|||||

134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150

|||||

389 TGATGTGATGAACGATTAATCATGTAACAGCGAATATAGACCCGTTATTC 438

|||||

151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167

|||||

439 AGAAATCTGAATATATAGCTCCGGCAGAGAGGATGTACGATTTACAGGTTT 488

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51326

seq_documentation_block:

ID AAQ51326 standard; DNA: 711 BP.

AC AAQ51326;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-114 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;


```

XX  ADP-ribosyltransferase activity; ss.
OS  Escherichia coli.
FH  Key Location/Qualifiers
FT  misc_difference 328..330
FT  /*tag=a
FT  /note="wild-type TCT(Ser) mutated to AAA(Lys)"
XX  PN
XX  WO9313202-A.
XX  PD
XX  08-JUL-1993.
XX  PF
XX  30-DEC-1992; 92MO-EP03016.
XX  PR
XX  31-DEC-1991; 91IT-0M13513.
XX  PA
XX  (BIOC-) BIOCLINE SCLAVO SPA.
PI  Domenighini M, Hol W, Pizsa M, Rappuoli R;
XX  WP1: 1993-227320/28.
XX  P-PSDB: AAR44025.
XX  DR
XX  P-PSDB: AAR44025.
XX  PT
XX  Immunogenic detoxified mutant cholera toxin and heat labile toxin
XX  enterotoxin producing Escherichia coli
XX  PT
XX  Claim 3; Fig 2 and Page 46; 60pp; English.
XX  PS
XX  The wild-type sequence coding for the A subunit of the heat labile
XX  toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
XX  et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
XX  site-directed mutagenesis. Certain mutations were found to reduce
XX  CC toxicity (see AA051314-Q051326). The invention relates to
XX  immunogenic, detoxified LT-A proteins and their use in vaccines to
XX  protect against enterotoxigenic E.coli. Sequence AA051326 is a
XX  combination of the wild-type coding sequence and the mutagenic
XX  primer sequence used to introduce the preferred mutation. (Amino
XX  acid numbering is based on the cholera toxin A subunit sequence).
XX  SQ
XX  Sequence 711 BP: 237 A; 122 C; 156 G; 196 T; 0 other;

alignment_scores:
    Quality: 1179.00      Length: 240
    Ratio: 5.171          Gaps: 2
    Percent Similarity: 95.000    Percent Identity: 92.917

alignment_block:
US-09-528-682-1 x AA051326 ..

Align seg 1/1 to: AA051326 from: 1 to: 711

1  ASnGLYAspArgLeuTYrARGLAAspSerARgPRoPAspGLuileLY 17
1  |||||
1  AATGGGACAGATTTATACCGTGTCTACGCTGAGACCCCAAGATGAATAA 50
1  |||||
1  sARgSerGLyGLyLeuMetPRoARgGLYHIsAsnGLyUTYrPhAspArgg 34
1  |||||
1  AGCTTTCGCGAGCTTATGCCAGAGGT...AATGAGTACTTTCGATAGAG 97
1  |||||
34  LYThGLMetAsnLLeAsnLeuTYrAspHISALARGLYThGLIoPhr 50
1  |||||
98  GAACTCAAATGAAATTTAATCTTTATGTATCCAGCGAGGAAACACAAACC 147
1  |||||
51  glyPheVALArgTYrAspAspGLYTYrVALSerThrSerLeuSerLeuAR 67
1  |||||
148  GGCCTTTCAGATATGATGACGAGATATGTTCACACTTCTCTAGTTTGAG 197
1  |||||
67  gserALAhISLeuAlaGLyGLnSerLLeuSerGLYTYrSerThrTYrTT 84
1  |||||
198  AAGTGTCTACTTATGACGACGACATATATTTATTCAGATATTCCTACTACTA 247

```

seq_name	seq	seq_name	seq	
84	YLLTFYValllElarhrralPProAsmMePhAsnValAsnspval	100	101	LeuglyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG1
248	TATATATCTGTTATAGCA.....AATATCTTTAAATGTTAATATGTA	288	289	ATTAGGATATACAGCCCTCCACCATATATGAACAGAGGTTAAAGCGTTAGG
117	YGLYlIleProTyrSerGlnlIleTyrGlyTrpTyrArvAlaSnPheGlyY	134	134	alllAspGluArGLeuHISArGAsnArGluTyrArGspArGTrpTyrTyr
339	TGGAATACCATATCTTCAGATATATGAGATGATCGTGTAAATTTTGGTG	388	389	TGATTTGATGAACGATTTACATCTGTAACAGGAAATATAGAACCCGGATATTC
151	ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArGLeuAlaGlyPh	167	167	ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArGLeuAlaGlyPh
439	AAAAATCTGAATATAGCTCCGGCAGAGGATGGTTACAGATTAGCAGGTTT	488	439	AAAAATCTGAATATAGCTCCGGCAGAGGATGGTTACAGATTAGCAGGTTT
167	eProProAspHisGlnAlaTrpArGluGluProTrpIleHISHisAlaP	184	167	eProProAspHisGlnAlaTrpArGluGluProTrpIleHISHisAlaP
489	CCCACCGGATTCACCAAGCTTGGAGAGAAACCCCTGATTCATCTGCAC	538	489	CCCACCGGATTCACCAAGCTTGGAGAGAAACCCCTGATTCATCTGCAC
184	roGlnGlyCysGlyAsnSerSerArGTrpIleThrGlyAspPhrCysAsn	200	184	roGlnGlyCysGlyAsnSerSerArGTrpIleThrGlyAspPhrCysAsn
539	CACAAGGTTTGAGAGTTCATCATCAAGACATATACAGGTCATCTTGTAAT	588	539	CACAAGGTTTGAGAGTTCATCATCAAGACATATACAGGTCATCTTGTAAT
201	GIuGIuTrpGlnAsnLeuSerThrIleTyrLeuArGluTyrGlnSerly	217	201	GIuGIuTrpGlnAsnLeuSerThrIleTyrLeuArGluTyrGlnSerly
589	GAGGAGACCCAGATCTGAGCACAATATATCTCAGGAAATATCAATCAAA	638	589	GAGGAGACCCAGATCTGAGCACAATATATCTCAGGAAATATCAATCAAA
217	SValysArGlnIlePheSerAspTyrGlnSerGluValaAspIleTyrA	234	217	SValysArGlnIlePheSerAspTyrGlnSerGluValaAspIleTyrA
639	AGTTAAAGAGCAGATATTTTCAGACTATCATCGACAGAGTTGACATATATA	688	639	AGTTAAAGAGCAGATATTTTCAGACTATCATCGACAGAGTTGACATATATA
234	snArGilleArGAspGluLeu	240	234	snArGilleArGAspGluLeu
689	ACAGAAATTCGGATGATTA	708	689	ACAGAAATTCGGATGATTA
seq_name	/SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051316			
seq_documentation_block	ID AA051316 standard; DNA; 711 bp.			
AC	AA051316;			
XX				
XX	08-DEC-1993 (first entry)			
DE	Encodes Tyr-53 E.coli heat labile toxin subunit A.			
XX				
XX	enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A.			
KM	Proctomer A; site-directed mutagenesis; reduced toxicity;			
XX	ADP-ribosyltransferase activity; ss.			
XX				
OS	Escherichia coli.			
XX				
XX	key	location/Qualifiers		
FT	misc_difference 153..155			
FT	FT	/tag= a		
FT	FT	/note= "wild-type GTC(Val) mutated to TAC(Tyr)"		
XX				
PN	W09313202-A.			
XX				
PD	08-JUL-1993.			
XX				
PF	30-DEC-1992; 92WO-EP03016.			
XX				
PR	31-DEC-1991; 91IT-OM13513.			
XX				
XX	(BIOC-) BIOCLINE SCLAVO SPA.			

CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SQ Sequence 711 BP; 235 A; 124 C; 155 G; 197 T; 0 other;

alignment_scores:
 Quality: 1178.00 Length: 240
 Ratio: 5.189 Gaps: 2
 Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:
 US-09-528-682-1 x AA051319 ..

Align seg 1/1 to: AA051319 from: 1 to: 711

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
1 AATGGCGACAGATTATACCGCTGACTAGACCCCGAGTGAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 ACGTTTCCGAGCTTATGCCAGAGGT...AATGAGTACTTCGATGAG 97
34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
98 GAACCTCAATGATATATTAATCTTATGATCAGCGAGAGGACCAACAC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
148 GCGTTTGTCAATATGATGACGAGTATGTTCCACTCTCTAGTTTGA 197
67 gSerAlHisIleuAlaGlnSerIleLeuSerGlyTyrSerThrTyrT 84
198 AAGTCTCAGTACGACAGACGATATATTCAGATATTCAGTTACTTA 247
84 YrlIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATAGCA.....AATATGTTTAATTACAATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
289 AATTGCGCTATACAGCCCTCAGATGACAGGAGGTTCTCGGTTAGG 338
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
339 TGGATATACCATATCTCAGATATATGATGATGATCGTTRATTTTGTG 388
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
389 TGATTTGATGAACGATTCATCGTAAACAGGAAATATAGAACCCGTTAT 438
151 ArgAsnIleuAsnIleAlaProAlaGlnAspGlyTyrArgLeuAlaGly 167
439 AGAATCTGATATATAGCTCCGCGAGAGATGTTACAGATTAGACAGTT 488
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisHisAla 184
489 CCCACCGGATCACCAAGCTTCGAGAGAGAAGAACCTTGATTCATCATG 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAGGTTGTGAGATTTCATCAAGAACATCAACAGTGAATCTGTAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrIleuArgGluTyrGlnSer 217
589 GAGAGAGACCCAGAACTTGAGACAAATATATCTCAGGAAATATCAAT 638
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
639 AGTTAAAGGACAGATATTTTCAGACTTCAGTCAAGAGTTGACATATATA 688
234 snArgIleArgAspGluLeu 240

```

|||||
 689 ACAGAAATTCGAGATGAATTA 708

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051315

seq_documentation_block:
 ID AA051315 standard; DNA; 711 BP.

AC AA051315;

XX 08-DEC-1993 (first entry)

DE Encodes Glu-53 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

FT Key Location/Qualifiers

FT misc_difference 153..155

FT /*tag= a

PD 08-JUL-1993.

PF 30-DEC-1992; 92MO-EP03016.

PR 31-DEC-1991; 91IT-0M13513.

PA (BIOC-) BIOGINE SCLAVO SPA.

PI Domenighini M, Hol W, Pizsa M, Rappuoli R;

DR WPI; 1993-227320/28.

DR P-PSDB; AAR38731.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

PS Claim 3; Fig 2 and page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile

CC toxin (LT-A) of a strain of E.coli known to affect humans (Tiamoto

CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to

CC site-directed mutagenesis. Certain mutations were found to reduce

CC toxicity (see AA051314-051326). The invention relates to

CC immunogenic, detoxified LT-A proteins and their use in vaccines to

CC protect against enterotoxigenic E.coli. Sequence AA051315 is a

CC combination of the wild-type coding sequence and the mutagenic

CC primer sequence used to introduce the preferred mutation. (Amino

CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 236 A; 122 C; 156 G; 197 T; 0 other;

alignment_scores:
 Quality: 1177.00 Length: 240
 Ratio: 5.185 Gaps: 2
 Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:
 US-09-528-682-1 x AA051315 ..

Align seg 1/1 to: AA051315 from: 1 to: 711

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
1 AATGGCGACAGATTATACCGCTGACTAGACCCCGAGTGAATAA 50

```



```

248 TATATATCGTTATAGCA.....AATATGTTTAAATAGATGATGTA 288
101 LeuGlyValIyrSerProHISProTYrGluGlnGluValSerIleLeuG 117
289 ATTAGGCTATACAGCCCTCCACCATATGACAGAGAGGTTCTCGTTAGG 338
117 yGlyIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPhelGly 134
339 TGGAAATACCATATTCAGATATATGATGATCGTATCGTTAATTTGGTC 388
134 aIleAspGluArgLeuHISArgAsnArgGluTYrArgAspArgTYrTYr 150
389 TGAATGTGATGACGATTACATCGTACACGGGAATATACAGCCGTTATAC 438
151 ArgAsnLeuAsnIleAlaProIleGluAspGlyTYrArgLeuIleGlyPh 167
439 AGAAATCTGAATATAGCTCCGGCAGAGATGTTACGATTACAGAGTTT 488
167 eProProAspHisGlnAlaIleAlaIleProArgGluGluProTYrIleHISAla 184
489 CCCACCGGATCACCAAGCTTGAGAGAGAAACCTGGATTCAATCATGACAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAAGGTTGTGAGATTTCATCAAGAACATCACAGGTGATCTGTAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSerI 217
589 GAGAGACCCAGAAATCGACACATATATCTCAGGGAATATCATGATAA 638
217 sValysArgGlnIlePheSerAspTYrGlnSerGluValAspIleTYr 234
639 AGTTAAAGGACAGATATTTTCACACTATCATGACAGGTTGACATATATA 688
234 snArgIleArgAspGluLeu 240
689 ACAGAAATTCGGAGTGAATTA 708

seq_name: /SID5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AA051314
seq_documentation block:
ID AA051314 standard: DNA; 711 BP.
XX
AC AA051314;
XX
XX 08-DEC-1993 (first entry)
DE Encodes Asp-53 E.coli heat labile toxin subunit A.
XX
XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KW ADP-ribosyltransferase activity; ss.
XX
XX Escherichia coli.
XX
XX Key Location/Qualifiers
XX FT misc_difference 153..155
XX FT /tag= a
XX FT /note= "wild-type GTC(Val) mutated to GAT(Asp)"
XX
XX W09313202-A.
XX
XX PD 08-JUL-1993.
XX
XX PF 30-DEC-1992; 92MO-EP03016.
XX
XX PR 31-DEC-1991; 91IT-OM13513.
XX
XX PA (BIOC-) BIOCINE SCLAVO SPA.
XX
XX PI Domenighini M, Hol W, Pizza M, Rappuoli R;
XX
XX WPI; 1993-227320/28.

```

```

DR P-PSDB; AAR38730.
XX
XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PI enterotoxin producing Escherichia coli
XX
XX
XX Claim 3; Fig 2 and Page 46; 60pp; English.
XX
XX The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051314 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (amino
CC acid numbering is based on the cholera toxin A subunit sequence).
XX
XX Sequence 711 BP; 235 A; 122 C; 156 G; 198 T; 0 other;

alignment_scores:
Quality: 1176.00 Length: 240
Ratio: 5.181 Gaps: 2
Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:
US-09-528-682-1 x AA051314 ..

Align seg 1/1 to: AA051314 from: 1 to: 711
1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIleuY 17
1 AATGGGACAGATTATACCGCTGCTGACTAGACCCCGAGATAAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTYrPhasArgG 34
51 ACCTTTCCGGAGTCTTATGCCAGAGGT...AATGAGTACTTGATAGAG 97
34 IyThrGlnMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrGlnI 50
98 GAACCTCAATGAATATTAATCTTATGATCAGCGAGAGAACACAAACC 147
51 GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeu 67
148 GCGTTTGATGATATGATGATGCGATATGTTCCACTCTCTGTTGAG 197
67 gSerAlaHisLeuAlaGlyIleSerIleLeuSerGlyTYrSerThrTY 84
198 AAGTGTCACTTAGCAGACAGTATATATATATCAGGATATTCACCTTA 247
84 yIleTYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATAGCA.....AATATGTTTAAATGATATGATGTA 288
101 LeuGlyValIyrSerProHISProTYrGluGlnGluValSerIleLeuG 117
289 ATTAGGCTATACAGCCCTCCACCATATGACAGAGAGGTTCTCGTTAGG 338
117 yGlyIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPhelGly 134
339 TGGAAATACCATATTCAGATATATGATGATCGTATCGTTAATTTGGTC 388
134 aIleAspGluArgLeuHISArgAsnArgGluTYrArgAspArgTYrTYr 150
389 TGAATGTGATGACGATTACATCGTACACGGGAATATACAGCCGTTATAC 438
151 ArgAsnLeuAsnIleAlaProIleGluAspGlyTYrArgLeuIleGlyPh 167
439 AGAAATCTGAATATAGCTCCGGCAGAGATGTTACGATTACAGAGTTT 488
167 eProProAspHisGlnAlaIleAlaIleProArgGluGluProTYrIleHISAla 184

```



```

|||||
489 CCCACCGGATCACCAGCTTGAGAGAGAACCTGGATTCAATCAGCAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAAGGTGTGGATTCATCAAGAACATCAACAGCTGACTTGTAAAT 588
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIly 217
589 GAGGAGACCCGAGATCTGAGCACAATATATCTCAGGAATATCAATCAAA 638
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
639 AGTTAAGAGGACAGATATTTTCACACATATCATCAGTCAAGCTTGACATATA 688
234 snArgIleArgAspGluLeu 240
689 ACAGAAATTCGGGATGAATTA 708

```

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51320

seq_documentation_block:
ID AAQ51320 standard: DNA; 711 BP.

AAQ51320;

08-DEC-1993 (first entry)

Encodes Gln-107 E.coli heat labile toxin subunit A.

enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

protomer A; site-directed mutagenesis; reduced toxicity;

ADP-ribosyltransferase activity; ss.

Escherichia coli.

Key Location/Qualifiers

misc_difference 307..309

/*tag=

/note= "wild-type CAC(His) mutated to GAG(Glu)"

WO9313202-A.

08-JUL-1993.

31-DEC-1991; 91IT-0M13513.

(BIOC-) BIOCINE SCLAVO SPA.

Domenighini M, Hol W, Pizsa M, Rappuoli R;

WPI, 1993-227320/28.

P-PSDB; AAR44019.

Immunogenic detoxified mutant cholera toxin and heat labile toxin

- useful as vaccines against infection by Vibrio cholerae and

enterotoxin producing Escherichia coli

Claim 3; Fig 2 and Page 46; 60pp; English.

The wild-type sequence coding for the A subunit of the heat labile toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto et al, J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to site-directed mutagenesis. Certain mutations were found to reduce toxicity (see AAQ51314-051326). The invention relates to immunogenic, detoxified LT-A proteins and their use in vaccines to protect against enterotoxigenic E.coli. Sequence AAQ51320 is a combination of the wild-type coding sequence and the mutagenic primer sequence used to introduce the preferred mutation. (Amino acid numbering is based on the cholera toxin A subunit sequence).

sq Sequence 711 BP; 234 A; 121 C; 158 G; 198 T; 0 other;

alignment_scores:

Quality: 1175.00 Length: 240
Ratio: 5.154 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 92.917

alignment_block:

US-09-528-682-1 x AAQ51320 ..

Align seg 1/1 to: AAQ51320 from: 1 to: 711

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIly 17
1 AATGGCGACAGATTATACCGTCTGACTGTAGACCCCAATGATAAATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 ACCTTCCGAGCTCTATGCCAGAGT...AATGACTACTTGATAGAG 97
34 IyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
98 GAACCTCAATGAATATATATCTTATGATCAGCGAGAGGACACAAACC 147
51 GlyPheValArgTyrAspArgGlyTyrValSerThrSerLeuSerLeu 67
148 GGCTTGTTCAGATATGATGAGGATATGTTCCACTCTCTGTAGTTGAG 197
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
198 AAGTGCCTCATTCAGCAGACAGATATATATATCAGATATTCATCTTA 247
84 yrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
248 TATATATCGTTATATAGCA.....AATATGTTAATGTTAATGATGTA 288
101 LeuGlyValIySerProHisProTyrGlnGlnValSerAlaLeuG 117
289 AATGAGCTATACAGCCCTGAGCCATATGAAACAGAGGATTTCTCGGTGAG 338
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrValAsnProGly 134
339 TGGATATCCATATTCAGATATATGATGATGATCGTGAATTTTGCTG 388
134 aIlLeAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyr 150
389 TGAATGATGAACGATTACATCGTAACAGGAAATATAGACCCGTAATAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
439 AGAAATCTGAATATAGCTCCGGCAGAGAGATGTTACGATTTACAGGTTT 488
167 eProProAspHisGlnAlaThrArgGluGluProTyrIleHisAla 184
489 CCCACCGGATCACCAGCTTGAGAGAGAACCCCTGGATTCAATCAGCAC 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
539 CACAAGGTGTGGAGATTCAACAAAGAACATCAACAGCTGACTTGTAAAT 588
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIly 217
589 GAGGAGACCCGAGATCTGAGCACAATATATCTCAGGAATATCAATCAAA 638
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
639 AGTTAAGAGGACAGATATTTTCACACATATCATCAGTCAAGCTTGACATATA 688
234 snArgIleArgAspGluLeu 240
689 ACAGAAATTCGGGATGAATTA 708

```


seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AA051324

seq_documentation_block:

ID AA051324 standard; DNA: 711 BP.

XX AA051324:

XX 08-DEC-1993 (first entry)

DE Encodes Ser-106 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

XX Key Location/Qualifiers

XX mis_c_difference 304..306

XX FT /note="wild-type CCT(Pro) mutated to AGT(Ser)"

XX MO313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92MO-EP03016.

XX 31-DEC-1991; 91IT-0MT3513.

XX (BIOC-) BIOICINE SCLAVO SPA.

XX Domenghini M, Hol W, Pizza M, Rappuoli R;

XX WPI: 1993-227320/28.

XX P-PSDB: AAR44023.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX PT - useful as vaccines against infection by Vibrio cholerae and

XX PT enterotoxin producing Escherichia coli

XX PS Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile

XX toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto

XX et al. J Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to

XX CC site-directed mutagenesis. Certain mutations were found to reduce

XX CC toxicity (see AA051314-Q51326). The invention relates to

XX CC immunogenic, detoxified LT-A proteins and their use in vaccines to

XX CC protect against enterotoxigenic E.coli. Sequence AA051324 is a

XX CC combination of the wild-type coding sequence and the mutagenic

XX CC primer sequence used to introduce the preferred mutation. (Amino

XX CC acid numbering is based on the cholera toxin A subunit sequence).

XX SQ Sequence 711 BP; 235 A; 122 C; 157 G; 197 T; 0 other;

alignment_scores:

Quality: 1175.00 Length: 240

Ratio: 5.176 Gaps: 2

Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:

US-09-528-682-1 x AA051324

Align seg 1/1 to: AA051324 from: 1 to: 711

1 AsnglyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIy 17

1 AATGGGACAGATTAACCGTCTGACCTAGACCCCGATGAATAA 50

17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34

51 ACCTTCCGGAGCTTATGCCAGAGT...AATGAGTACTTCGATACAG 97

34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50

98 GAACCTCAATGATATTAATCTTTATGATCAGCGAGGAACACAAAC 147

51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67

148 GGCCTTGTCAGATATGATGACGAGATATGTTCCACTTCCTTACGTTGAG 197

67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84

198 AAGTCTCACTTAGCAGACAGATATATATATATCAGAGATATCTACTTA 247

84 YrIleTyrValIleAlaThrIleProAsnMetPheAsnValAsnAspVal 100

248 TATATATCGTTATACGA.....AATATCTTAATGTTATATGATGA 288

101 LeuGlyValTyrSerProHisProTyrGlnGlnIleValSerAlaLeuG 117

289 ATTACGATATACAGCAGCCCATATGACAGAGGTTTCGCTTACG 338

117 YGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134

339 TGGAAATACCATATCTCAGATATATGATGATGATGTTAATTTGGTG 388

134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150

389 TGATTCATGACAGATTAATCTTAACAGGAGATATAGAGACCGGATATAC 438

151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167

439 ACMAATCTGATATATAGCTCCGACAGAGATGCTTACAGATTAGCGAGTT 488

167 eProProAspHisGlnAlaThrArgGluGluProThrIleHisAlaP 184

489 CCCACCGGATCACCAAGCTTGAGAGAGAACACCTGGATATATACGAC 538

184 rGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200

539 CACAAGTGTGGAGATTCATCATACAGACATACAGAGTGTATCTGTAT 588

201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217

589 GAGGAGACCCAGCAATCTGACCAATATATCTCAGGAGATATCAATCAA 638

217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234

639 AGTTAAGAGCGAGATATTTTCAGACTATATCAGAGGTTGACATATATA 688

234 sNArgIleArgAspGluLeu 240

689 ACAGAAATTCGGATGATTA 708

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AA051321

seq_documentation_block:

ID AA051321 standard; DNA: 711 BP.

XX AA051321:

XX 08-DEC-1993 (first entry)

DE Encodes Lys-104 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

XX ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

XX Key Location/Qualifiers

XX mis_c_difference 298..300

PT - useful as vaccines against infection by *Vibrio cholerae* and
 PT enterotoxin producing *Escherichia coli*

PS Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of *E. coli* known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic *E. coli*. Sequence AA051323 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 234 A; 122 C; 157 G; 198 T; 0 other;

alignment_scores:

Quality: 1174.00 Length: 240
 Ratio: 5.172 Gaps: 2
 Percent Similarity: 94.583 Percent Identity: 92.917

alignment_block:

US-09-528-682-1 x AA051323 ..

Align seg 1/1 to: AA051323 from: 1 to: 711

```

1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProAspGluIleTy 17
1 AATGGCAGCAGATATATACCGTGCCTAGACCCCAAGATGAAATAAA 50
17 sATgSerGlyGlyLeuMetProArgGlyHisAsnGluTyRPhesArg 34
51 ACGTTTCGGAGCTTATGCCAGAGGT...AATGACTTCTTCATRGAG 97
34 lYthRImetAsnIleAsnLeuTyRAspHisAlaArgGlyThRGlThR 50
98 GAACCTCAATGAATATATATCTTATGATCAGCGAGAGACAAAC 147
51 GlyPheValArgTyRAspArgGlyTyValSerThRserLeuArg 67
148 GCGTTTGTCAATATGATGAGCGATATGTTCCACTCTTGATTGAG 197
67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTyRserThR 84
198 AAGGCTCAGCTTAGCAGACAGATATATATCAGATATTCACCTTA 247
84 yRlIeTyValIleAlaThRAlaProAsnMetPheAsnValAsnAsp 100
248 TATATATCGTTATATGCA.....AATATGTTTAAATGATGATGA 288
101 LeuGlyValIlySerProHisProTyRGlUGlUGlValSerAlaLeu 117
289 ATTAAGCGTAAGTACCCCTCAACATATGAACAGAGGTTTCTGCGT 338
117 yGlyIleProTySerGlnIleTyRGlTyRPyRArgValAsnPhes 134
339 TGGAAATACCATATTCAGATATATGATGATGATCGTTAAATTTG 388
134 allIeAspGluArgLeuHisArgAsnArgGluTyRArgAspArg 150
389 TGAATGTAGAACGATATCATCGTAACAGGAAATATAGAACCGGTA 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyRArgLeuAla 167
439 AGAATGTGATATATAGCTCCGCAAGAGATGTATACGATTAGACG 488
167 eProAspHisGlnAlaIleThRArgGluGluProThRIleHisIle 184
489 CCCACCGGATACCAAGCTTGGAGAGAGAACCCGTGATTCATCATC 538

```

```

184 roGlnGlyCysGlyAsnSerSerArgThRIleThRGlYAspThRCysAsn 200
539 CACAAGCTGTGGAGATCATCAAGAACATCAACAGCTGATCTGTATAT 588
201 GlUGlUthRGlnAsnLeuSerThRIleTyRleuArgGluTyRGlnSer 217
589 GAGAGACCCAGCAATCTGAGACAAATATATCTCAGGAAATATCAACAA 638
217 sValIysArgGlnIlePheSerAspTyRInsSerGluValAspIleTy 234
639 AGTTAAGAGGACATATTTTCAGACTATATCAGTCAGAGCTTGACATATA 688
234 sNArgIleArgAspGluLeu 240
689 ACAGAAATTCGGGATGAATTA 708

```

seq_name: /SIDS5/gcdata/geneseq/geneseqn_emb1/NA1993.DAT:AA051322

seq_documentation_block:

ID AA051322 standard; DNA; 711 BP.

AC AA051322;

DT 08-DEC-1993 (first entry)

DE Encodes Asp-104 *E. coli* heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS *Escherichia coli*.

FH Key Location/Qualifiers

FT misc_difference 298..300 /*tag= a

FT /*note= "wild-type TAC(Tyr) mutated to GAT(asp)"

XX WO9313202-A.

XX PN 08-JUL-1993.

XX PD 30-DEC-1992; 92WO-EP03016.

XX PF 31-DEC-1991; 91IT-0M13513.

XX PR (BIOC-) BIOCINE SCLAVO SPA.

XX PA Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX PT WPI; 1993-227320/28.

XX DR P-PSDB; AAR44021.

XX PS Claim 3; Fig 2 and Page 46; 60pp; English.

CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of *E. coli* known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic *E. coli*. Sequence AA051322 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 234 A; 122 C; 157 G; 198 T; 0 other;

84 yrlletyValIILlelAhhriaAProasmePheasnValasnapVal 100
251 ATATATATATGTTATACCCACTGCACCCCAACATGTTAACTTATGATGTA 300
101 LeuGIyValItyrSerProHIsprOtyrGIuGIInglValSerAlaLeuGI 117
301 TTAGGGCGATACAGTCCCTCATCCACATGACAGAAAGTTTCTGCTTTAGG 350
117 yGIyIleProtyrSerGIInIleTyrgIyTPtyrAArgValasnPheGIyV 134
351 TGGGATTTCCATACGCCCAATATATATGATGATGATGATGATTTGGGG 400
134 alIleaspluarGIleuHIsarGaanAArgIuTyTyrAArgspargIyTyTyr 150
401 TGCCTTATGAAACAATATCATGCTAATAGCGGTACAGACATAGATATTAC 450
151 ArgasnlleuasnIleAlaProIaGIuaspGIyTyTyrArgIleuAlaGIyph 167
451 AGTAACTTAGATATTGCTCCACACACAGATGGTATTGATGGATGGCAAGTTT 500
167 eProProaspHIsGlnAlaIATrPArGIuGIuProTrpIleHIsHsAlaAP 184
501 CCCGCCGAGCATAGAGCTTGGAGGGAAGAGCGCTGATTCATCATGAC 550
184 roGIInGIyCyGIasnserserserargthrIleTherGIyaspHrCysasn 200
551 CGCGGGGTGTGGGAATCTCCCAACATCATTCGATCATGTAATACCTTCGAT 600
201 GluGIuThrGIasnleuSerThrIleTyTyrleuAArgIuTyrcInserIy 217
601 GAAAAAACCAAAGCTTAGGTATAAATTCCTTGCAGCAATACCAATCTAA 650
217 sValIysArgIInIlePheSerAspTyrcInserGIuValasIleTyTyrA 234
651 AGTAAAGACAAATATTTTCAGCGCTATCAATCTGATATGATGATACACATA 700
234 snArgIleArGAsplueu 240
701 ATAGAAATTAAGATGAAATTA 720

```

seq_name: /SDBS5/gcgdata/geneseq/geneseqn-embL/NA2000.DAT:AAA51157
seq_documentation_block:
ID   AAA51157 standard;  CDNA, 777 BP.
XX
AC   AAA51157;
XX
DT   26-SEP-2000   (first entry)
XX
DE   Plant-optimized V. cholerae toxin A subunit coding sequence.
XX
KW   Heat-labile toxin; CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
KM   adjutant; anti-bacterial; ss.
XX
OS   Vibrio cholerae.
XX
OS   Synthetic.
XX
FH   Key
CD   Location/Qualifiers
FT   1..777
FT   /*tag= a
FT   /product= cholera_toxin_subunit_A
XX
PN   WO200037609-A2.
PD   29-JUN-2000.
XX
XX   22-DEC-1999;   99WO-US30747.
XX
XX   22-DEC-1998;   98US-0113507.
XX
XX   (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA   (MASON/) MASON H S.
PA   (ARNTZEN/) ARNTZEN C J.

```

XX Mason HS, Arintzen CJ;
PI
XX WPI: 2000-442653/38.
DR P-PSDB; AA976653.
XX
XX New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX Disclosure; Fig 4A; 103pp; English.

This plant/codon optimized cDNA encodes a synthetic Vibrio cholerae
CC cholera toxin (CT) A subunit (CT-A). The sequence has no cryptic signal
CC sequence and no CG (potential methylation sites) sequences. Novel
CC polynucleotides encode a mutant Escherichia coli heat-labile toxin (LT)
CC subunit A (LT-A) polypeptide or a mutant V. cholerae cholera toxin (CT)
CC A subunit (CT-A) polypeptide, which have reduced enzyme activity as
CC compared to the wild-type LT-A or CT-A polypeptide and where at least one
CC of the codons is altered to a plant preferred codon. The polynucleotide
CC further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or
CC a CT B subunit (CT-B). The polynucleotides are useful for the
CC transformation of plant cells for the production of transgenic plants to
CC produce edible vaccines, especially oral vaccines in transgenic plants
CC for the prophylactic or therapeutic treatment against E. coli or
CC V. cholerae. The mutant polypeptides are also useful as adjuvants.

SO Sequence 777 BP; 209 A; 174 C; 191 G; 203 T; 0 other;

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x AAA51157 ..

Align seg 1/1 to: AAA51157 from: 1 to: 777

1 ASngGLVAsPArGLeuTYrArgAlaAsPSeRarPRoProAsPgLuIllely 17
||| |||:::|||||
55 AAdGaTCAAGAcTCTATAgGGcACACTCAgACCtCCTGTGAATCAA 104

17 sArSergLyGlYdEumetPrOArGclYnHAsngLntyrPhesArPg 34
:::|||||
105 GCATCATGAGTGCGTCTATGCCAAAGGGACAATCTGACTTTGCACAGG 154

34 LyrHrgImetAsnIleasnLeuTyrrsrnHisAlaarGlyThncLthr 50
|||||
155 GTACTCGATGACATCAACCTTTATACCCATGCAAGGGCAACTCAACT 204

51 GlYPheValArGTyTAAsPaSPglTYrValSerThrSerLeuSerLeuAr 67
|||||
205 GGAATTGTGAGGCGATGATGATGATGTGTGCACACTCCATTAGCTTAG 254

67 gSerAlaHisLeuAlaclYginSerTlleuSerGlyTyISerThryTT 84
|||||
255 GTCTGCCACACTTGTGGTGCACAACTACCTCTCTGCGTCACTACTACT 304

84 YrlleTyValIleAlaThraLarProasmetPheaSnValAsnaSraVal 100
|||||
305 ACATCTATGTGATGGCCATGSCATGCCAACATGTTCAATGTATAATGATGTG 354

101 LeuglYValtyrSerProHisrToTyrgLUngInluvalSerAlalaLeuGI 117
|||:::|||||
355 TTGGAGCACAAACAGCCCTCCACAGAGACAGCAAGAAGGTCTGCTTGGG 404

117 uGllleArtoTySerGlnlinleTyrgLutTrTyrgAlaValasnrheguY 134
|||||
405 TGGAATCCCACTACCTCCAATCTATGGATGGATTAAGGTGTGACATTGGAG 454

134 allleArsgluArGleuHisAlgrAsnArGrGluTYrATrGAsPaRArGTyTr 150


```

455 TCCTTGATAGGCAACTCCATAGGAATAGGGCGTACAGGATAGGTACTAC 504
|||||:::||||||| |||||||
151 ArgAsnLeuAsnIleAlaProAlaLysPglYtyrArgLeuAlaGlyPh 167
|||||:::||||||| |||||||
505 AGCAACTTGACATTGCCTCCACACAGCAGATGTTATGGATTGGCAGGTTT 554
|||||:::||||||| |||||||
167 eProFroAspHisGlnAlaTrpArgLuglIProTrpIleHisIslaP 184
|||||:::||||||| |||||||
555 CCCTCAGACAGCATAGGCGCTGGAGGAGAGACCTTGGATTTCACCATTGCAC 604
|||||:::||||||| |||||||
184 roGLngLYCYsgLYasNserSerArgrhrIlethrGlyAspThrCYsaSn 200
|| |:::||||||| ||:::|
605 CACCAAGGTGTGTGAAATCTCCCAAGTCACAGCATGACCAACAACCTGTGAT 654
|||||:::||||||| ||:::|
201 gUGuLUThrgInAsnLeuSerThrIleTyrlEuArGsuLUryrGlnserLy 217
|||||:::||||||| ||:::|
655 GAAAAGACCCAAATCTTGGGTGTGAAGTCTCTGATGAGTACCACTTA 704
|||||:::||||||| ||:::|
217 sValLYsrArgGlnIlePheSerAspTYrGlnSerGluValAspIleTyra 234
|||||:::||||||| ||:::|
705 GGTTGAAGAGGCCAANCTTCTCAGCGCTACCACATTCATTCACACCCACA 754
|||||:::||||||| ||:::|
234 snArGtlLEArGAspLcUleu 240
|||||:::||||||| ||:::|
755 ATAGATCTCAGGATGAACTC 774

seq_name= /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: AAN30169
seq_documentation_block:
ID AAN30169 standard; DNA: 1148 BP.
XX
XX AAN30169;
XX
XX 25-APR-1992 (first entry)
XX
XX Sequence of PRfR 18014 encoding sub-unit A and B of cholera toxin.
XX
XX Vaccine; cholera; subunit A; subunit B; ss.
XX OS Vibrio cholerae.
XX
XX Key Location/Qualifiers
FH 1..777
FT misc_feature /*tag= a
FT /label= F1
FT /note= "claimed fragment"
FT misc_feature 774..1148
FT /*tag= b
FT /label= F2
FT /note= "claimed fragment"
XX EP95452-A.
XX PN
XX 30-NOV-1983.
XX PD
XX PE 23-MAY-1983; 83EP-0091416.
XX PF
XX PR 24-MAY-1982; 82US-0381083.
XX PA (SMIK ) SMITH KLINE-RIT.
XX PI
XX Harford N, Dewilde M;
XX DR WPI: 1983-834665/49.
XX
XX DNA sequence coding for cholera toxin sub units - for
XX PT transfecting host cells, useful in vaccine prodn.
XX PS
XX Claim 2; Page 27-28; 46pp; French.
XX
XX The inventors claim a sequence conty. at least one fragment coding for
XX CC for all or part of sub-units A and B of cholera toxin. The entire

```

```

CC sequence and two fragments (F1 and F2) are listed in the
CC specification. Also new are recombinant DNA molecules contg. these
CC sequences (spec. PR1T 18014 contg. the whole sequence; PR1T 18041
CC contg. F1 and PR1T 10810 contg. F2) and transformed host cells (spec
CC E. coli ATCC 39052 contg. the whole sequence, ATCC 39053 contg. F1
CC and ATCC 39051 contg. F2).
XX
XX
50 Sequence 1148 BP; 382 A; 184 C; 235 G; 347 T; 0 other:

Alignment_scores:
    Quality: 1088.00      Length: 240
    Ratio: 4.814          Gaps: 0
    Percent Similarity: 94.167      Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x AAN30169 ...

Align seg 1/1 to: AAN30169 from: 1 to: 1148

1  AasnGlyAspArgLeuTyrIrrgAlaAspSerArgProPoaSpgLileTy 17
   ||| |||:::|||||
55  AATGATGATATAGTATATATGCGGAGAGTTCTGAGCTCTGTATGTAATAA 104
   |||
17  SAISerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArgC 34
   :::::
105  GCACTCGAGTGCGCTTATGCAAGAGCAGAGTCACTCTTGACGACG 154
   |||
34  LyrTrpGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
   |||
155  GTACTCAAAATGAATATCAACCTTTATGATCATGCAAGAGCAACGACG 204
   |||
51  GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuArg 67
   |||:::
205  GGATTTGTAGCGACGATGATGATATGTTCCACTCAATATGATTGAG 254
   |||
67  gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
   |||:::
235  AAGTGCCCACTTGTAGGGGTCAAACATATGTCGGTATATCTTATTT 304
   |||
84  YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
   |||
305  ATATATATGTTATATAGCCACGACGCCAACATGTTTAACGTTAATGATGTA 354
   |||
101  LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuGln 117
   |||:::
355  TTAGGGGCATACAGTCCCTCATCCAGATGACAGAAAGTTTCTGCTTAGG 404
   |||
117  YGlyIleProTyrSerGlnIleTyrGluTyrPyrTyrValAsnGlyIle 134
   |||:::
405  TGGGATTCATCTACTCCCAATATATGATGATGATGAGATCTATTTTGGGG 454
   |||
134  aIleAspGluArgLeuHisIleArgAsnArgGluTyrArgAspArgTyrTyr 150
   :::::
455  TGCCTTGATGAACAATTACATCGTAATAGGGGCTACAGACGATAGATATTAC 504
   |||
151  ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlnPhe 167
   |||:::
505  AGTAACTAGTATATGCTCCAGCAGCAGTGGTATGGATTGGCAGAGTTT 554
   |||
167  eProPAspHisGlnAlaAlaTrpArgGluGluIleProTrpIleHisHisAlaP 184
   |||:::
555  CCGTCGCGAGCATAGAGCTTGGAGGGAAGAACGCCGTGATCATCATGCAC 604
   |||
184  rGlnGlnGlyCysGlyAsnSerSerArgThrIlePheGlyAspThrCysAsn 200
   |||:::
605  CGCGCGGCTGTGGGAATGCTCCCAAGATCATCATGAGTAATATCTTGGAT 654
   |||
201  GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIle 217
   |||:::
655  GAAATAAACCCAAAGTCTAGGTGTAAATATCTCTTGACGATATCCATATGAA 704
   |||
217  sValIysArgGlnIlePheSerAspTyrIleSerGlnValAspIleTyrA 234

```



```

|||||
705 AGTTAAAGACAAATTTTTCAGGCTATCAATCTGATATGTACACATA 754
234 snArg1leArgAspGluLeu 240
755 ATAGAAATTAAGATGAATTA 774

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA167270

seq_documentation_block:
ID AA167270 standard; DNA: 2056 BP.
XX
XX AA167270:
XX
XX
XX 11-FEB-2002 (first entry)
XX
XX Cholera toxin A and B subunits encoding DNA sequence.
XX
XX Enterotoxin; mutation; chimeric; immunostimulant; immune response;
XX Cholera toxin; ds.
XX
XX Vibrio cholerae.
XX
XX Key Location/Qualifiers
XX 5'UTR 1..550
XX /*tag- a
XX CDS 551..1327
XX /*tag- b
XX /*product- "Cholera toxin A subunit"
XX /*note- "AAG65991"
XX CDS 1324..1698
XX /*tag- c
XX /*product- "Cholera toxin B subunit"
XX /*note- "AAG65992"
XX /*product- "B chain"
XX 3'UTR 1699..2056
XX /*tag- d
XX
XX NO200170257-A1.
XX
XX 27-SEP-2001.
XX
XX 16-MAR-2001; 2001MO-US08582.
XX
XX 17-MAR-2000; 2000US-190058P.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Mcghee J, Kiyono H, Takeda Y, Ohmura M, Yamamoto S;
XX
XX WPI: 2001-648368/74.
XX P-PSDB: AAG65991, AAG65992.
XX
XX Chimeric molecule useful as adjuvant for cell-mediated/humoral immunity
XX PT comprises first mutated A subunit of first enterotoxin and second
XX PT non-mutated subunit from second enterotoxin different from natural
XX PT enterotoxin -
XX
XX Disclosure: Fig 4; 23pp; English.
XX
XX PS The invention provides a chimeric molecule that comprises a first subunit
XX CC which is mutated A subunit of first enterotoxin and a second non-mutated
XX CC subunit from a second enterotoxin which is different from the natural
XX CC enterotoxin which has been mutated to provide A subunit. A composition of
XX CC matter comprising the chimeric molecule in a pharmaceutically acceptable
XX CC carrier is useful for obtaining enhanced immune response of an organism
XX CC to an antigen by administering it to the antigen. The chimeric molecule
XX CC is useful to provide specific immune response to a particular
XX CC enterotoxin, as adjuvants for use with unrelated vaccines, and to
XX CC customize adjuvants to direct production of cell-mediated or humoral
XX CC immune responses. It is also useful as mucosal adjuvant for cell-mediated
XX CC or humoral immunity. The present sequence represents the DNA sequence of
XX CC Cholera toxin A and B subunits.

```

```

XX
SQ Sequence 2056 BP; 603 A; 327 C; 479 G; 647 T; 0 other;

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x AA167270 ..

Align seg 1/1 to: AA167270 from: 1 to: 2056

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluLeu 17
|||
605 AATGATGATAGTAAGTATATCGGCAGATTCCTGATGAAATATA 654
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
|:::|
655 GCAGTCAGGTGGCTTATGTCGACAGAGACAGTGAATGATGACCGAG 704
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
705 GTACTCAATGATATCAACCTTATGATCATGCAAGAGAACTGAGAG 754
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
|||||
755 GGATTTGTAGGCACGATGATGATGATGTTCCACCTCAATGATTTGAG 804
67 gSerIleHisIleuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
|||||
805 AAGTGGCCACTAGTGGTCAAACTATATGTCGTGCTTACTTATTT 854
84 yIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
855 ATATATATGTTATAGCAGCTGACCAACATGTTTAACTTATATGATGA 904
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG1 117
|||||
905 TTAGGGGCATACAGTCCATCCAGATGACAAAGAAAGTTTCGCTTAGG 954
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGlyVal 134
|||||
955 TGGGATTCCTACTCTCCANATATATGATGATGATGATGATTTTGGGG 1004
134 aIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
|||||
1005 TGCTTGATGACAAATATACATCGTAATAGGGCGCTACAGATATGATTTAC 1054
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhe 167
|||||
1055 AGTAACTTAGATATGCTCCACAGCAGCATGGTTATGATTTGGCAGATT 1104
167 eProProAspHisGlnAlaThrArgGlnGluProTyrIleHisAlaPhe 184
|||||
1105 CCGTCCGAGCAATAGAGCTTGGAGGAGAGCCGCTGATTCATCATGACAC 1154
184 rGlnGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||
1155 CCGCGGGTGTGGGAATGCTCCAGATCATCGATGATGATATCTGCGAT 1204
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGlyTyrGlnSerIly 217
|||||
1205 GAAAAAACCAAGTCAAGTGAATAATTCCTTGACGAATACCAATCTAA 1254
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAlaPheIleTyr 234
|||||
1255 AGTTAAAGACAAATTTTTCAGGCTATCAATCTGATATGTACACATA 1304
234 snArg1leArgAspGluLeu 240
|||||
1305 ATAGAAATTAAGATGAATTA 1324

```


KM Hybrid; fusion; membrane translocation; binding region; HIV;
 KM Infection; toxin; steroid; hormone; monoclonal antibody; antigen;
 KM diphtheria; exotoxin; phenylketonuria; cholera; interleukin-IL-2;
 KM protease; epidermal growth factor; ricin; tetanus; hexosaminidase;
 KM Shiga-like toxin A; ligand; insulin; nuclease; ss.

OS Vibrio cholera.

XX Key Location/Qualifiers
 FT 1..774
 CDS /tag= a
 /label= CT-A1

PN WC0109871-A.

XX 11-JUL-1991.

PF 21-DEC-1990; 90MO-US07619.

PR 14-JUN-1990; 90US-0538276.

PR 22-DEC-1989; 89US-0456095.

XX (SERA-) SERAGEN INC.

XX Murphy JR.

XX WPI; 1991-222845/30.

XX P-PSDB; AAR13117.

PT Hybrid molecules for targeting chemical entity to cell - have
 PT membrane trans-locating and cell binding-regions and used to
 PT treat HIV infection, genetic enzyme-deficiency disorders etc.

PS Disclosure; Fig 5(1-2); 59pp; English.

XX Hybrid molecules are produced by covalently linking
 CC (1) a portion (A) of the binding domain of a cell-binding ligand,
 CC allowing binding of the mol. to an animal cell;
 CC (2) a portion (B) of a translocation domain of a protein able to
 CC translocate (C) across the cell cytoplasmic membrane, and
 CC and (3) a portion (C) which is to be introduced into the cell.
 CC (A) is derived from a steroid or polypeptide hormone, a single-chain
 CC analogue of a monoclonal antibody able to bind an antigen expressed
 CC on the cell surface, or a polypeptide toxin.
 CC (B) is derived from a toxin (e.g. diphtheria toxin or Pseudomonas
 CC exotoxin A).
 CC (A) may be derived from insulin, interleukins 2, 3 or 6 or
 CC epidermal growth factor.
 CC Suitable enzymes in (C) include cholera toxin (encoded by this
 CC sequence), ricin, tetanus toxin, hexosaminidase A, protease,
 CC nuclease, etc.
 CC Specified examples are CT-A/DT-B'/IL-2, STRA/DT-B'/IL-2,
 CC Ricin A/DT-B'/IL-2, HIVP-BP/DT-B'/IL-2 and the phenylalanine
 CC hydroxylase-DT-B' or their biologically active mutants.
 CC (CT-A= cholera toxin, DT-B'= truncated diphtheria toxin,
 CC STRA= Shiga-like toxin A; HIVP-BP= HIV protease binding protein.
 CC See also AAQ12710-12.

XX Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other;

XX Alignment_scores:

XX Quality: 1085.00 Length: 240
 XX Ratio: 4.801 Gaps: 0
 XX Percent Similarity: 94.167 Percent Identity: 81.250

XX Alignment_block:

XX US-09-528-682-1 x AAQ12710 ..

XX Align seg 1/1 to: AAQ12710 from: 1 to: 780

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluIleLeu 17
 ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

55 AATGATGATAGTATATATCGGCAGATTCTAGACCTCTGATGAATATA 104
 17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 105 GCAGTCAGGTGGTCTTATGCCAAGAGGACAGAGTACGATCTTGACCCGG 154
 34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 155 GTACTCAAAATGAAATATCAACCTTTATGATCATGCAAGAACTCAGACG 204
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
 205 GGATTTGTTAGGCACGATGATGATATGTTCCACCTCAATAGTTGAG 254
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
 255 AAGTCCCACTTAGTGGGTCAACATATATGCTGTCATCTACTATAT 304
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 305 ATATATATGTTATAGCCACTGCCACCCACATGTTTACCTTATGATGTA 354
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
 355 TTAGGGGCAATACAGTCCATCCAGATGAAACAAGAGTTTCGCTTAGG 404
 117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGlyV 134
 405 TCGGATTCATCATCTCCCAAAATATGAGATGATCGACATTTTGGGG 454
 134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 455 TCGTTGATGACAAATTCATCGTAATAGGGGCTACAGAGATGATATATAC 504
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
 505 AATACTTATGATATGCTCCACAGCAGCATGGTTATGATGGAGGTTT 554
 167 eProProAspHisGlnAlaTrpArgGluGluProTrpIleHisAlaP 184
 555 CCCTCGGAGCATAGAGCTTGAGGAGGAGAGCCGCGGATTCATCATGAC 604
 184 rGlnGlyValGlyAsnSerArgThrIleThrGlyAspThrCysAsn 200
 605 CCGCGGTTGTGGGAATGCTCCAAAGATCATCGATGATATCTTCCGAT 654
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerI 217
 655 GAAAAAACCAGTCTAGTGTAAATTCCTTGACGATACCAATCTAA 704
 217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
 705 ACTTTAAACACAAATATTTTCAGCGTATCAATCTGATTATGATACACA 754
 234 snArgIleArgAspGluLeu 240
 755 ATAGATTTAAGATGAAATTA 774

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AA91636

seq_documentation_block:

ID AAT91636 standard; DNA; 780 BP.

AC AAT91636;

DT 27-MAR-1998 (first entry)

DE Vibrio cholerae toxin gene.

KW Cholera toxin A1; hybrid protein; cell delivery;

KW cell binding ligand; translocation domain; diphtheria toxin B';

KW interleukin-2; leukemia; therapy; ss.

OS Vibrio cholerae.

XX Key Location/Qualifiers

FT CDS

1.777

FT /*tag= a

/transl_except= (pos:358..360, aa:Ala)

FT

/transl_except= (pos:616..618, aa:Ala)

XX

US566825-A.

XX

16-SEP-1997.

XX

07-JUN-1984.

XX

07-JUN-1984.

XX

27-JUN-1991.

XX

07-JUN-1984.

XX

25-APR-1985.

XX

07-JUN-1985.

XX

22-DEC-1989.

XX

14-JUN-1990.

XX

04-AUG-1993.

XX

93US-0102387.

XX

(SERA-) SERAGEN INC.

XX

Murphy JR.

XX

WPI: 1997-470103/43.

XX

P-PSDB: AAW25785.

XX

New hybrid molecules for delivery of agents to cells - comprise a

XX

binding domain of a cell binding ligand and a portion of a

XX

translocation domain of a protein

XX

Example 2; Fig 5A-B; 30pp: English.

XX

This sequence comprises the Vibrio cholerae gene that encodes

XX

cholera toxin (see AAW25785). DNA encoding an enzymatically active

XX

Al fragment of the cholera toxin was used to construct a cholera

XX

toxin A1 diphtheria toxin B'-interleukin-2 (CTA-DTB-IL2) gene that

XX

was expressed in E. coli. The hybrid protein can be isolated and

XX

used e.g. as an adjunct to treatment with DTB-IL2 hybrid, which

XX

targets the cell-killing ability of diphtheria toxin to cells

XX

bearing the IL2 receptor, such as certain leukaemic T cells.

XX

Prior application of the CTA-DTB-IL2 hybrid alleviates the

XX

problem of endogenous IL2 competing with DTB-IL2 hybrid for IL2

XX

receptors on T cells. Claimed hybrid proteins comprise a

XX

translocation domain and a cell binding domain from e.g. a hormone,

XX

growth factor or polypeptide toxin. The hybrid molecules can be

XX

used for the delivery of agents (e.g. therapeutic genes, toxins,

XX

detectable labels) into cells. The use of a translocation

XX

mechanism ensures that the hybrid will be effective in relatively

XX

low doses, since a high proportion of the substance of interest

XX

will be taken into the targeted cells. The hybrid molecules can be

XX

manufactured as a single hybrid recombinant protein, permitting

XX

CC reproducibility, consistency, and the precise control of

XX

composition.

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

XX

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661

seq_name: /SIO5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAZ30661


```
XX OS Vibrio cholerae.
XX PN US9565406-A.
XX PD 12-OCT-1999.
XX PF 07-JUN-1995; 95US-0488246.
XX PR 04-AUG-1993; 93US-0102387.
PR 07-JUN-1994; 84US-0618199.
PR 27-JUN-1991; 91US-0722484.
PR 25-APR-1985; 85US-0726808.
PR 07-JUN-1985; 85US-0742554.
PR 22-DEC-1989; 89US-0456095.
PR 14-JUN-1990; 90US-0538276.
XX (SERA-) SERAGEN INC.
XX PI Murphy JR.
XX DR WPI: 1999-632431/54.
XX DR P-PSDB: AAY55890.
XX PT Recombinant DNA molecule encoding a three part hybrid protein used in
PT the treatment of Aids and genetic deficiency diseases -
XX PS Example 2; Fig 5; 31pp: English.
XX CC The invention relates to a recombinant DNA molecule encoding a hybrid
CC protein comprising three parts: (a) the first part comprises a portion
CC of the binding domain of a cell-binding polypeptide ligand allowing the
CC hybrid protein to bind to an animal cell; (b) the second part comprises
CC a portion of a translocation domain of a naturally occurring protein
CC selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera
CC toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus
CC toxin, which translocate the third part of the across the cytoplasmic
CC membrane into the cytosol of the cell; and (c) the third part comprises
CC a polypeptide entity to be introduced into the cell, which is non-native
CC to the naturally occurring protein of (b). This sequence represents the
CC Vibrio cholerae toxin A1 fragment gene for use in generating the hybrid
CC of the invention. The hybrid molecule enables the direction of
CC appropriate therapy to affected cells, allowing them to function properly
CC and alleviate or cure the disease. The hybrid is especially used in
CC treating genetic deficiency diseases, by delivering to affected cells
CC an enzyme supplying the missing function, to supplementing cellular
CC levels of a particular enzyme or a scarce precursor or cofactor, to
CC directing toxins or other poisons to destroy particular cells (such as
CC adipocytes, cancer cell, or virus infected-cells), to counteracting viral
CC infections such as HIV, by introducing appropriate antibodies to viral
CC substances. It is also involved in the process of getting non-therapeutic
CC substances such as detectable labels into cells.
XX Sequence 780 BP: 243 A: 130 C: 169 G: 238 T: 0 other:
XX
```

```
alignment_scores:
    Quality: 1085.00      Length: 240
    Ratio: 4.801          Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.250
```

```
alignment_block:
US-09-528-682-1 x AA230661 ..
```

```
Align seg 1/1 to: AA230661 from: 1 to: 780
```

```
1 AsnG1AspArgLeuTyrArgAlaAspSerArgProProAspGluIleu 17
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
55 AATGATGATTAAGTATATCGGCGCATTCAGACCTCCGATGATAATATA 104
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
105 GCAGTCAGGTCGCTTATGCGCAAGAGACAGACTAGTACTTGACCGAG 154
```

```
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
155 GTACTCAAAATGAATATCAACCTTTATGATCATGCAAGAGAACTCAGAC 204
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
51 GlyPheValArgTyrAspArgGlyTyrValSerThrSerLeuSerLeu 67
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
205 GGATTTGTTAGCGACATGATGATGATGTTTCCACCTCAATTAGTTGAG 254
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
255 AAGTCCCACTAGTGGGTCACAACTATATGCTGGCATTCACCTATAT 304
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
84 YrlIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnArg 100
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
305 ATATATATGTTATAGCCACATGACCCCAACATGTTTAACTTAATGAT 354
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
355 TTAGGGGATACAGTCTCTCATCCAGATGAACAGAACTTCTCTTAGG 404
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
117 YglYlIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheG 134
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
405 TGGGATTCCTACTCCCAATATATGATGATGATGATTCATTTGGGG 454
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
134 allLeasprGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
455 TGCTTGATGAACAATTACATCAATAGGGGCTCAAGAGATGATATTC 504
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
151 ArgAsnLeuAsnIleLeuAlaProAlaGluAspGlyTyrArgLeuAla 167
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
505 AGTAAGTTCATGATATGCTCCAGCAGCAGATGTTATGATGGCAGTTT 554
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
167 eProProAspHisGlnAlaThrArgGluGluProTyrIleHisAlaAl 184
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
555 CCTCCGGACATAGAGCTTGAGGAGAGCCGTGATTCATCATGACAC 604
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
184 roGlnGlySerGlyAsnSerSerArgThrIleThrGlyAspThrCysAs 200
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
605 CGCGGGTGTGGGAATGCTCCAAAGATCAATGATGATGATTCGAT 654
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
201 GluGlnThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
655 GAAAAAACCAAGCTAGGTGTAATTCCTTGACGAATACCAATCTTA 704
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
705 AGTTAAAGCAATATATTTTCAGCGTATCAATCTGATATGATACACATA 754
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
234 snArgIleArgAspGluLeu 240
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
755 ATGAATTTAAGATGAATTA 774
```

```
seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AA290017
```

```
seq_documentation_block:
ID AA290017 standard: DNA: 780 BP.
```

```
AA290017:
```

```
05-MAY-2000 (first entry)
```

```
Vibrio cholerae toxin gene.
```

```
Cholera toxin: hybrid protein; translocation domain; cell binding domain;
genetic deficiency disease; cell targeting; cell destruction; cancer;
adipocyte; enzyme delivery; anti-viral; HIV, ds.
```

```
Vibrio cholerae.
```

```
US6022950-A.
```

```
08-FEB-2000.
```



```

XX 08-JUL-1993.
XX 30-DEC-1992: 92WO-EP03016.
XX 31-DEC-1991: 91IT-0MI3513.
XX (BIOC-) BIOGINE SCLAVO SPA.
XX Domenighini M, Hol W, Pizsa M, Rappuoli R;
XX WPI; 1993-227320/28.
XX P-PSDB; AAR44027.
XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
XX - useful as vaccines against infection by Vibrio cholerae and
XX enterotoxin producing Escherichia coli
XX
XX Claim 3: Fig 2 and Page 48; 60pp; English.
XX
XX The wild-type sequence coding for the A subunit of the cholera
XX toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
XX was subjected to site-directed mutagenesis. Certain mutations were
XX found to reduce toxicity (see AA051327-051334). The invention relates
XX to immunogenic, detoxified CT-A proteins and their use in vaccines
XX to protect against Vibrio cholerae. Sequence AA051328 is a
XX combination of the wild-type coding sequence and the mutagenic
XX primer sequence used to introduce the preferred mutation.
XX
XX Sequence 723 BP; 228 A; 124 C; 164 G; 207 T; 0 other;

```

```

alignment_scores:
    Quality: 1084.00      Length: 240
    Ratio: 4.796          Gaps: 0
    Percent Similarity: 94.167    Percent Identity: 81.250

```

```

alignment_block:
US-09-528-682-1 x AA051328 ..

```

```

Align seg 1/1 to: AA051328 from: 1 to: 723

```

```

1 AasnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluLeu 17
111 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 AATGATGATTAAGTATATCGGCGAGATTCTAGACCTCGATGAAATTA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
111:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 GCAGTCAGGCGCTTATGTCACAGAGACAGACTGATCTTGACGAG 100
34 LyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 GTAAGTCAAAATGAATATCAACCTTATGATCATGCAAGAGAACTCAG 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 GGAATTGTTAGCGACGATGATGATGTTTCCACCAAGATTGTTGAG 200
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 AAGTGCCCACTTATGTCGTCAAACCTATATGTCGTCATTTCTACCT 250
84 yTlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 ATATATATGTTATAGCCACCTGACCCACACATGTTTAAACGTTAAT 300
101 LeuGlyValTyrSerProHisProTyrGlnGluValSerAlaLeuGln 117
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 TTATGGGCGATACAGTCCTCATCCAGATGAACAAGATTCTCTTAGG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 TGGGATTCCATCTCCCAATATATATGATGATGATGATTCATTGGGG 400

```

```

134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
111:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 TGCTTGATGACAAATTATCATCTGTAATAGGGGCTACAGAGATTAAT 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
451 AGTAACTTATGATTTGCTCCAGCAGACATGCTTATGATTTGGCAGTT 500
167 eProProAspHisGlnAlaThrArgGluGluProTrrIleHisAlaAla 184
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
501 CCTCCGAGCATATAGACTTGGAGGAGAGACCGCTGATTCATCATGCA 550
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
551 CGCCGGGTGTGGAAATGCTCCAAAGATCATGATCATGATTAATCTGC 600
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
111:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601 GAAAAAACCCAAAGTCTAGTGTAATAATCTTGACGAATACCAATCTA 650
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
651 AGTTAAAGACAATATTTTACAGCTATCATCTGATTAATGATACACAT 700
234 snArgIleArgAspGluLeu 240
111|||||:|||||:|||||:|||||:|||||:|||||:|||||:
701 ATGAAATTAAGATGAATTA 720

```

```

seq_name: /SID5/gcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051334

```

```

seq_documentation_block:

```

```

ID AA051334 standard; DNA: 723 BP.

```

```

AC AA051334;

```

```

DT 08-DEC-1993 (first entry)

```

```

DE Encodes Glu-114 cholera toxin subunit A.

```

```

XX Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

```

```

KW Protomer A; site-directed mutagenesis; reduced toxicity;

```

```

KW ADP-ribosyltransferase activity; ss.

```

```

OS Vibrio cholerae.

```

```

XX Key Location/Qualifiers

```

```

FH misc_difference 340..342 /tag= a

```

```

FT /note= "wild-type TCT(Ser) mutated to GAA(Glu)"

```

```

XX WO9313202-A.

```

```

XX 08-JUL-1993.

```

```

XX 30-DEC-1992: 92WO-EP03016.

```

```

XX PR 31-DEC-1991: 91IT-0MI3513.

```

```

XX PA (BIOC-) BIOGINE SCLAVO SPA.

```

```

XX PI Domenighini M, Hol W, Pizsa M, Rappuoli R;

```

```

XX WPI; 1993-227320/28.

```

```

XX DR P-PSDB; AAR44033.

```

```

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

```

```

XX - useful as vaccines against infection by Vibrio cholerae and

```

```

XX enterotoxin producing Escherichia coli

```

```

XX Claim 3: Fig 2 and Page 48; 60pp; English.

```

```

XX The wild-type sequence coding for the A subunit of the cholera

```



```

AC AA051332:
DT 08-DEC-1993 (first entry)
DE Encodes Ser-110 cholera toxin subunit A.
XX Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;
KM protomer A; site-directed mutagenesis; reduced toxicity;
XX ADP-ribosyltransferase activity; ss.
OS Vibrio cholerae.
XX
XX key location/Qualifiers
XX misc_difference 328..330
XX FT /*tag a
XX FT /note= "wild-type GAA(Glu) mutated to TCG(Ser)"
XX
XX W09J313202-A.
XX
XX 08-JUL-1993.
XX
XX 30-DEC-1992; 92MO-EP03016.
XX
XX 31-DEC-1991; 91IT-OMI3513.
XX
XX (BIOC-) BIOCIINE SCLAVO SPA.
XX
XX Domenighini M, Hol W, Piazza M, Rappuoli R:
PI
XX WPI: 1993-227320/28.
XX DR
XX P-PSDB: AAR44031.
XX
XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli
XX
XX Claim 3; Fig 2 and Page 48; 60pp; English.
XX
XX The wild-type sequence coding for the A subunit of the cholera
CC toxin (CT-A) (Nekalanos et al., Nature 306, 551 (1983) - see AA046318,
CC was subjected to site-directed mutagenesis. Certain mutations were
CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051332 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
XX Sequence 723 BP; 225 A; 126 C; 163 G; 209 T; 0 other;
XX
XX Alignment_scores:
XX Quality: 1083.00 Length: 240
XX Ratio: 4.792 Gaps: 0
XX Percent Similarity: 94.167 Percent Identity: 81.250
XX
XX Alignment block:
XX US-09-528-682-1 x AA051332 ..
XX
XX Align seg 1/1 to: AA051332 from: 1 to: 723
XX
XX 1 AsnGlyAspArgLeuTyrFArgAlaAspSerArgProProAspGluIlely 17
XX ||| |||:::|||||
XX 1 AATGATGAATGAATTATATCGGCGAATTCATGACCTCTCGATGAAATAAA 50
XX
XX 17 SARGSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArgG 34
XX ||:::|||||
XX 51 GCAGTCAGAGTGTCTTATGCCAAGAGCAGACAGAGTACTTGTGACCGAG 100
XX
XX 34 IYThGImeCAsnIleAsnLeuTyrAspHisAlaArgGlyThnGlnThr 50
XX |||||
XX 101 GTACGCAAAATGAAATATCAACTTTATGATCATGCAAGAGACTCAGACG 150
XX
XX 51 glyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67

```


seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NN1993.DAT.AA051333

seq_documentation_block:

ID AA051333 standard; DNA; 723 BP.

XX
AC AA051333;

XX DT 08-DEC-1993 (first entry)

XX DE Encodes Ala-112 cholera toxin subunit A.

XX KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;
KW Protomer A; site-directed mutagenesis; reduced toxicity;
KW AP-ribosyltransferase activity; ss.
XX OS Vibrio cholerae.

XX FH Key Location/Qualifiers
FH misc_difference 334..336
FT /*tag= a
FT FT /note= "Wild-type GAA(Glu) mutated to GCT(Ala)"
XX PN K09J313202-A.
XX PD 08-JUL-1993.
XX PF 30-DEC-1992; 92MO-EP03016.
XX PR 31-DEC-1991; 91IT-OMI313.
XX PA (BIOC-) BIOICINE SCLAVO SPA.
XX PI Domenighini M, Hol W, Pizzi R, Rappunni R;
PI WPI, 1993-227320/28.
XX DR P-PSDB; AAR4032.
XX PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli
XX PS Claim 3; Fig 2 and Page 48; 60pp; English.

The wild-type sequence coding for the A subunit of the cholera toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318) was subjected to site-directed mutagenesis. Certain mutations were found to reduce toxicity (see AA051327-051334). The invention relates

CC	to immunogenic, detoxified Cr-A proteins and their use in vaccines
CC	to protect against Vibrio cholerae. Sequence AAQ51333 is a
CC	combination of the wild-type coding sequence and the mutagenic
CC	primer sequence used to introduce the preferred mutation.
xx	
50	Sequence 723 BP; 225 A; 126 C; 163 G; 209 T; 0 other;
	alignment_scores:
	Quality: 1082.00 Length: 240
	Ratio: 4.809 Gaps: 0
	Percent Similarity: 93.750 Percent Identity: 81.250
	alignment_block:
	US-09-528-682-1 x AAQ51333 ..
	Align seg 1/1 to: AAQ51333 From: 1 to: 723
1	AsnGlyAspArgLeuTyrTrgAlaAspSerAlaGpProPdsGluIlely 17
1	
1	AAAGGATGATTAAGTTATATCGGAGATTCTTAACCTCCGATGAATAATAA 50
17	sARSGELyGLyLeuMetProARGLyHisnGLyTrpPheAspArg 34
17	
51	GCAGTAGAGGTGGCTTATGCCAAGAGAGACAGATGATCTTGGACCGAG 100
34	lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgSlyThrGlnThr 50
101	GTACTCAATGATATCAACCTTTATGATCATGCAAGAGAACTGAGACG 150
51	GlyPheValArgTyrAspSglyTyrValSerThrSerLeuSerLeuAr 67
151	GGATTTGTTAGCCACAGCATGATGATATGTTCCACCTCAATAGATTAG 200
67	gSerAlaHisIleuAlaGlyGlnSerIleLeuSerSlyThrThrTyrT 84
201	AAATGGCCACTTAGTGGGTCAAACATATATGCTGCTGATCTTACTGTT 250
84	yrIleTyrAlaIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251	ATATATATGTTATAGCCACTGCACCCCAACATGTTTAACTTATATATGTA 300
101	LeuGlyValTyrSerProHisProTyrGluGlnIleValSerAlaLeuGl 117
301	TTAGGGCCATACAGTCCCTCATGCAGATGACAACGTTGTTCTGCTTAGG 350
117	yGlyIleProTyrSerGlnIleTyrGlyTrpTyrArgValAsnPheGlyY 134
351	TGGGATTTCCATACCTCCCAATATATGGAGTGGATTCAGATCTATTTGGGG 400
134	alleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
401	TGCTTATATACACATTTACATCTGTAATAGGGGCTACAGAGATATGATATTC 450
151	ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
451	AGTAACCTTATGATATTCCTCCACAGACAGATGGTTATGATTCGACGTTT 500
167	eProPdsPheHisGlnAlaTrpArgGluGluProTrpIleHisAlaIap 184
501	CCCTCCGGAGCATAGAGCTTGGAGGGAAGACCGCTGGATTCATTCATGAC 550
184	roGlnIyGysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
551	CGCCGGGTTGGGAATGCTCCAAAGTATCATCATCAATATATCTTGCGCAT 600
201	gluGluThrGlnAsnLeuSerThrIleTyrLeuArgGlyTyrGlnSerly 217
601	GAATAAACCCAAAGTCTAGCTATAAATTCCTTGAAGAAATACCATCTA 650
217	sVallySarGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
651	AGTTAAAGACAAATATTTTCAGGCTATCAATCAATCATATTTGATACACATA 700

234 snArglleArgAspGluLeu 240
 |||||:::|||||
 701 ATAGATTAAAGATCAATTAA 720

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT.AA051327

seq_documentation_block:

ID AA051327 standard; DNA: 723 BP.

XX AA051327:

DT 08-DEC-1993 (first entry)

DE Encodes Asp-53 cholera toxin subunit A.

KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Vibrio cholerae.

FT Key Location/Qualifiers

FT misc_difference 156..158

FT /*tag= "a" /note= "wild-type GTT(Val) mutated to GAC(Asp)"

PN MO9313202-A.

PD 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-0MT3513.

PA (BIOC-) BIOCTINE SCLAVO SPA.

PI Domenighini M, Hol W, Piazza M, Rappuoli R;

DR WPI: 1993-227320/28.

DR P-PSDB: AAR44026.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli

PS Claim 3; Fig 2 and Page 48; 60pp; English.

XX The wild-type sequence coding for the A subunit of the cholera
 CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
 CC was subjected to site-directed mutagenesis. Certain mutations were
 CC found to reduce toxicity (see AA051327-051334). The invention relates
 CC to immunogenic, detoxified CT-A proteins and their use in vaccines
 CC to protect against Vibrio cholerae. Sequence AA051327 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation.

XX Sequence 723 BP; 228 A; 126 C; 163 G; 206 T; 0 other;

alignment_scores:

Quality: 1081.00 Length: 240

Ratio: 4.804 Gaps: 0

Percent Similarity: 93.750 Percent Identity: 81.250

alignment_block:

US-09-528-682-1 x AA051327 ..

Align seg 1/1 to: AA051327 from: 1 to: 723

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
 ||| |||||:::|||||
 1 AATGATGATTAAGATTATATCGGCGCATCTAGACCTCGATGATAATAA 50

17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
 |:::|||||:::|||||
 51 GCAGTCAGGTGCTTATGCCAAGAGGACAGTACGATCTTGACCGAG 100
 34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
 |||||:::|||||
 101 GTATCTCAATGAATATCAACCTTATGATCATGCAAGAGGAACTCAGACG 150
 51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
 |||||:::|||||
 151 GGATTTGACAGCAGCATGATGATATGTTCCACCTCAATGATTTGAG 200
 67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrT 84
 |||||:::|||||
 201 AAGTCCCACTTATGAGGTCGCAACATATATGCTGCTGCTTACTTATTT 250
 84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
 |||||:::|||||
 251 ATATATATGTTATAGCCACTGCACCCACATGTTAAACGTTAATGATGA 300
 101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
 |||||:::|||||
 301 TTAGGGCATACAGTCCTCATCAAGATGACAGAGATTTCGCTTTAGG 350
 117 YGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGlyV 134
 |||||:::|||||
 351 TGGGATTCATCTCCCAATATATGATGATGATGATGATGATTTTGGGG 400
 134 aIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
 |||||:::|||||
 401 TCGTTGATGACAAATATACATCTAATAGGGCTACAGATGATGATATATAC 450
 151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
 |||||:::|||||
 451 AGTAACTTATGATATTTGCTCAGCAGCAGATGCTTATGATTTGGCAGTTT 500
 167 eProProAspHisGlnAlaTyrPArgGluGluProTyrPheHisAlaP 184
 |||||:::|||||
 501 CCTCCGGAGCATACAGTCTTGAGGGAAGAGCCGAGGATTCATCATGCAC 550
 184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
 |||||:::|||||
 551 CCCCGGTTGTGGGAATGCTCCCAAGATCATGATGATATCTGCGCAT 600
 201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
 |||||:::|||||
 601 GAAAAAACCAGTCTAGGTGAATAATCTTGACGAATACCAATCTTA 650
 217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
 |||||:::|||||
 651 AGTTAAACAAATATTTTTCAGGCTATCATCTGATATGATACACANA 700
 234 snArglleArgAspGluLeu 240
 |||||:::|||||
 701 ATAGATTAAAGATCAATTAA 720

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT.AA051331

seq_documentation_block:

ID AA051331 standard; DNA: 723 BP.

XX AA051331:

DT 08-DEC-1993 (first entry)

DE Encodes Asn-107 cholera toxin subunit A.

KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Vibrio cholerae.

SQ Sequence 777 BP; 208 A; 174 C; 191 G; 203 T; 1 U; 0 other;

Quality:	1075.00	Length:	240
Ratio:	4.799	Gaps:	0
Percent Similarity:	93.333	Percent Identity:	80.833

US-09-528-682-1 x AAA51160 .

Align seg 1/1 to: AAA51160 from: 1 to: 777

[illegible]

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAQ31263

ID ' AAQ31263 standard; DNA; 777 BP.

DT 25-MAR-1993 (first entry)

DE Mutant A subunit of CT.

OS *Vibrio cholerae*

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT slg-peptide

FT	mat_peptide	55..774
----	-------------	---------

misc_feature

misc_feature

800-678-9222

ET 2444010

ET
ET
ET

misc differe

ET
ET
misc differa

misc differe

misc differe

misc feature

XX 1

PD 12-NOV-1992.

XX

PA (AMGE-) AMGE

PI Burnette WN,

DR P-PSDB; AAR2

XX Recombinant DNA encoding analogue of Vibrio cholerae catalytic
PT sub-unit - useful as a vaccine for preventing cholera, has little
PT or no catalytic activity

PS Claim 7-8: Page 49-50 + Fig 1A; 74pp; English.

XX A recombinant mol. encoding an analogue of the catalytic subunit of
CC cholera toxin (CT), where the analogue has reduced or no catalytic
CC activity associated with CT reactivity is an analogue of the A
CC region, esp. A1 region, comprising a site-specific mutation in the
CC region bounded by the codons for Met1 and Arg192 or Ser194, esp. one
CC or more of Arg1, Arg11, Asp9, His44, His70 and Glu112 (see tag h-m
CC of Features Table). The analogue may also comprise a truncation of
CC the C-terminal portion starting at Trp179 (see tag n of Features, Table).
CC The DNA may also encode subunit B of CT, esp. the B oligomer in
CC native form or which has been genetically engineered (see AA031264).

XX Sequence 777 BP; 235 A; 125 C; 165 G; 234 T; 18 other;

alignment_scores:

Quality: 1045.00 Length: 240
Ratio: 4.750 Gaps: 0
Percent Similarity: 91.667 Percent Identity: 79.167

alignment_block:

US-09-528-682-1 x AA031263 ..

Align seg 1/1 to: AA031263 from: 1 to: 777

```

1  AAsGlyAspArgLeuTyArgAlaAspSerArgProProAspGluIleTy 17
   ||| ||||| ||| ||| ||||| ||||| |||
55 AATGATGATAGTATATATNNNGCANNNTCTCTGATGAAATATAA 104
   ||| ||||| ||||| ||||| ||||| ||||| |||
17 SAISerGlyGlyLeuMetProArgGlyHisAsnGluTyThrPheAspArg 34
   ||| ||||| ||||| ||||| ||||| ||||| |||
105 GCAGTCAGGTGGCTTATGCCAAGACAGAGTACTTGTACGACGAG 154
   ||| ||||| ||||| ||||| ||||| ||||| |||
34  LyThrGlnMetAsnIleAsnLeuTyArgPheHisAlaArgGlyThrGlnThr 50
   ||| ||||| ||||| ||||| ||||| ||||| |||
155 GTATCAAAAGTAATATCACTTATATGATNNNGCAAGAGGACCAAGC 204
   ||| ||||| ||||| ||||| ||||| ||||| |||
51  GlyPheValArgTyArgAspGlyTyArgValSerThrSerLeuSerLeuArg 67
   ||| ||||| ||||| ||||| ||||| ||||| |||
205 GGATTTGTAGGCACGATGATGATGTTTCCACCTCAATTAATGTTGAG 254
   ||| ||||| ||||| ||||| ||||| ||||| |||
67  gSerAlaHisLeuAlaGlyIleSerIleLeuSerGlyTySerThrTyT 84
   ||| ||||| ||||| ||||| ||||| ||||| |||
255 AAGGCGCANNNTTACTGGGTCAAACTATATGTCGTCATCTCTACTATT 304
   ||| ||||| ||||| ||||| ||||| ||||| |||
84  yTleTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
   ||| ||||| ||||| ||||| ||||| ||||| |||
305 ATAATATGTTATATAGCCACTGCACCAACATGTTTAACGTTAATGATGTA 354
   ||| ||||| ||||| ||||| ||||| ||||| |||
101 LeuGlyValTySerProHisProTyArgGluGlnValSerAlaLeuG 117
   ||| ||||| ||||| ||||| ||||| ||||| |||
355 TTAGGGGACATACAGTCTCCATCCAGATGAACAANNNGTTTCTGCTTAGG 404
   ||| ||||| ||||| ||||| ||||| ||||| |||
117 yGlyIleProTySerGlnIleTyArgTyPtyArgValAsnPheGly 134
   ||| ||||| ||||| ||||| ||||| ||||| |||
405 TGGGATTTCATATCCCAAAATATATGATGATATCGATCATTTTGGG 454
   ||| ||||| ||||| ||||| ||||| ||||| |||
134 alIleAspGluArgLeuHisArgAsnArgGluTyArgAspArgTyTyr 150
   ||| ||||| ||||| ||||| ||||| ||||| |||
455 TGCATTGTAACCAATTCATATGATAGGGGCTACAGAGATAGATATAC 504
   ||| ||||| ||||| ||||| ||||| ||||| |||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyArgGluAlaGly 167
   ||| ||||| ||||| ||||| ||||| ||||| |||
505 AGTAACCTAGATATTTGCTCAGCAGCAGATGTTATGATTTGCGAGTTT 554
   ||| ||||| ||||| ||||| ||||| ||||| |||
167 eProProAspHisGlnAlaThrParGluGluProTrpIleHisHisAla 184
   ||| ||||| ||||| ||||| ||||| ||||| |||

```

```

555 CCCTCCGAGCATAGACTTGGAGGAGAGCCGTCGATTCATCATGCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
   ||| ||||| ||||| ||||| ||||| ||||| |||
605 CGCCGGGTGTGGAAATGCTCCAAATCATCATGATGATTAATCTTCGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyArgGluTyArgGlnSerIle 217
   ||| ||||| ||||| ||||| ||||| ||||| |||
655 GAAAAAACCAAGTGTAGTAAATTCCTTGACGAATACCAATCACTAA 704
217 sValTyArgGlnIlePheSerAspTyArgGlnSerGluValAspIleTy 234
   ||| ||||| ||||| ||||| ||||| ||||| |||
705 AGTAAAGCAAAATATTTTCAGGCTATCAATCTGATATTTGATACACA 754
234 snArgIleArgAspGluLeu 240
   ||| ||||| ||||| ||||| ||||| ||||| |||
755 ATAGAAATTAAGATGATTA 774

```

seq_name: /SIDS5/gcdata/geneseq/geneseq-emb1/NA2002.DAT:ABA92644

seq_documentation_block:

```

ID  ABA92644 standard; DNA; 7586 BP.
XX
AC  ABA92644;
XX
DT  25-MAR-2002 (first entry)
XX
DE  Cholera toxin A1 subunit sequence SEQ ID NO:1.
XX
KW  Cholera toxin; bacterial bleb; vaccine; eukaryotic expression cassette;
   gene therapy; ds.
XX
Vn  Vibrio cholerae.
XX
WO  WO200109535-A1.
XX
PD  29-NOV-2001.
XX
PF  24-MAY-2001; 2001WO-US16904.
XX
PR  24-MAY-2000; 2000US-206994P.
XX
PA  (UYMA-) UNITV MARYLAND BIOTECHNOLOGY INST.
XX
PI  Powell RJ, Hone D;
XX
WP  2002-122002/16.
XX
PT  Introducing and expressing a gene in animal cells, useful for
   delivering a therapeutic agent to the cells, comprises infecting animal
   cells with bacterial blebs containing a eukaryotic expression cassette
   encoding the gene -
XX
PS  Example 5; Fig 3; 83pp; English.
XX

```

The present invention describes a method for introducing and expressing a gene in animal cells, comprising infecting animal cells with bacterial blebs containing a eukaryotic expression cassette encoding the gene.

The present invention also describes: (1) a bacterial bleb containing a eukaryotic expression cassette; (2) a transformed animal cell including a bacterial bleb, or a transfected product of the bacterial bleb; (3) a method of delivering a eukaryotic expression cassette to an animal cell, comprising mucosal transfection; and (4) a method of therapeutically treating an animal with a therapeutic agent encoded by a polynucleotide coding sequence, by transfecting the animal with a bacterial bleb containing a eukaryotic expression cassette including the polynucleotide coding sequence. The method can be used for introducing exogenous or foreign genes into animal cells using bacterial blebs as vectors. The method allows for the delivery of eukaryotic expression cassettes encoding the endogenous or foreign genes into animal cells or tissues, and is useful for expressing e.g. vaccine antigens, gene therapeutic agents, immunoregulatory agents, antisense RNAs, and catalytic RNAs in animal cells or tissues. The bacterial blebs

CC containing the eukaryotic expression cassette can also be used to treat
 CC animal cells cultured in vitro. The method can be used in gene therapy.
 CC The present sequence represents the cholera toxin A1 subunit sequence,
 CC which is used in an example from the present invention for the production
 CC of non-pyrogenic blebs containing a DNA vaccine.

Sequence 7586 BP; 1619 A; 1747 C; 1733 G; 1657 T; 830 other;

alignment_scores:
 Quality: 929.00 Length: 191
 Ratio: 5.049 Gaps: 0
 Percent Similarity: 96.335 Percent Identity: 86.911

alignment_block:
 US-09-528-682-1 x ABA92644 ..

Align seg 1/1 to: ABA92644 from: 1 to: 7586

```

3 AsparGLeuTrpArgAlaAspSerArgProProAspGluIleLysArgSe 19
|||||
9 GATAGCTTATCGGCGAGATCTTACACCTCGATGAATTAAGCAGTC 58
|||||
19 rglGlyLeuMetProArgGlyHisAsnGluTrpPheAspArgGlyThrG 36
|||||
59 AGGTGGCTTATGCCAAGAGCAGAGTGTACTTGAACCGAGTACTC 108
|||||
36 lMetAsnIleAsnLeuTrpAspHisAlaArgGlyThrGlnThrGlyPhe 52
|||||
109 AATGTGATATCAACCTTATGATCATGCAAGAGAACTCAGACGGGATT 158
|||||
53 ValArgTrpAspAspGlyTrpValSerThrSerLeuSerLeuArgSerAl 69
|||||
159 GTTAGGCACATGATGAGATGTTTCCACCTCAATTAAGTTGAGAAATGC 208
|||||
69 ahISleuAlaGlyInSerIleLeuSerGlyTrpSerThrTrpTyrIleT 86
|||||
209 CCACCTAGTGGTGCACAACTATATTGTCTGTCATCTACTATATATAT 258
|||||
86 yTAlaIleAlaThrAlaProAsnMetPheAsnValAsnAspValLeuGly 102
|||||
259 ATGTTATAGCCACTGCCACCATATGTTTAACTTATGATGATTAATGAGG 308
|||||
103 ValTyrSerProHisProTyrGluGlnGluValSerAlaLeuGlyIle 119
|||||
309 GCATTACAGTCTCATCCAGATGACAAAGATTTCTGCTTAGCGGGAT 358
|||||
119 eProTyrSerGlnIleTyrGlyTrpTyrArgValAsnPheGlyValIle 136
|||||
359 TCCATCTCCCAATATATGATGATGATGATGATGATGATGATGATGATG 408
|||||
136 spGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
409 ATGACAAATTAATCATGTAATAGGGCTACAGATAGATGATTAACAGTA 458
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProP 159
|||||
459 TTAGATATTGCTCCAGCAGCAGATGTTATGATGATGATGATGATGATG 508
|||||
169 cAspHisGlnAlaTrpArgGluGluTrpTrpIleHisAlaProGln 186
|||||
509 GGAGCATAGAGCTTGGAGGGAAGACCGTGTGATTCATCATGACACGGCG 558
|||||
186 lYcysGlyAsnSerSerArgThr 193
|||||
559 GTTGTGGGAATGCTCCAGATCA 581
|||||

```

seq_name: /SID55/gcgsdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAF74430

seq_documentation_block:

ID AAF74430 standard; DNA: 564 BP.

XX

AC AAF74430;

```

XX 08-MAY-2001 (first entry)
DT
XX
XX Vibrio cholerae ctxA gene 564 bp region nucleotide sequence.
DE
XX
XX Vibrio cholerae; ctxA gene; detection; identification; toxigenic; ds.
KW
XX
XX Vibrio cholerae.
OS
XX
XX ZA200001349-A.
XX
XX 27-DEC-2000.
XX
XX 16-MAR-2000; 2000ZA-0001349.
XX
XX 16-MAR-2000; 2000ZA-0001349.
XX
XX (UYPR-) UNIV PRETORIA.
XX
XX Theron J, Venter SN, Brozel VS, Du Preez M;
XX
XX WPI, 2001-191829/19.
XX
XX Novel oligonucleotide primer useful for PCR amplification of toxigenic
PT Vibrio cholerae nucleic acid for detecting toxigenic Vibrio cholerae
PT organisms present in the sample

```

Example 1; Fig 6; 44pp; English.

The present invention describes a method for amplifying (A) toxigenic
 CC Vibrio cholerae nucleic acid by: (a) performing a polymerase chain
 CC reaction (PCR), by combining a medium containing toxigenic V. cholerae
 CC nucleic acid with an oligonucleotide primer (I) which will bind to a
 CC section of V. cholerae nucleic acid, in an environment suitable for
 CC carrying out PCR; (b) combining the medium with a second primer which
 CC will bind downstream to the first primer to a section of the
 CC complementary strand of the V. cholerae nucleic acid; and (c) amplifying
 CC corresponding sections of V. cholerae nucleic acid present in the medium.
 CC The method is useful for detecting toxigenic V. cholerae organisms or
 CC toxigenic V. cholerae nucleic acid, e.g., ctxA gene, in an environmental
 CC water sample e.g., tap water sample, river water sample, lake water
 CC sample or a sewage sample. The method is useful in PCR assays for
 CC detecting a microorganism or nucleic acid of the microorganism in a
 CC sample. (i) and amplification products are useful for detecting toxigenic
 CC V. cholerae nucleic acid in a sample. The PCR method involving (I) is
 CC rapid, economical, specific and sensitive for detecting toxigenic
 CC V. cholerae in water samples. The present sequence represents a
 CC region from the V. cholerae ctxA gene, which is used in an example from
 CC the present invention.

Sequence 564 BP; 159 A; 105 C; 139 G; 161 T; 0 other;

alignment_scores:
 Quality: 910.00 Length: 187
 Ratio: 5.056 Gaps: 0
 Percent Similarity: 96.257 Percent Identity: 87.166

alignment_block:
 US-09-528-682-1 x AAF74430 ..

Align seg 1/1 to: AAF74430 from: 1 to: 564

```

7 ArgAlaAspSerArgProProAspGluIleLysArgSerGlyLysLeu 23
|||||
1 CGGGCAGATCTTAGACCTCTGATGAATTAACAGCAGTCAAGTGTCTTAT 50
|||||
23 tProArgGlyHisAsnGluTrpPheAspArgGlyThrGlnMetAsnIle 40
|||||
51 GCCAAGAGGACAGAGTGAAGTCTTGAACGAGGTACTCAATGAATATCA 100
|||||
40 snLeuTrpAspHisAlaArgGlyThrGlnThrGlyPheValArgTyrAsp 56
|||||

```



```
101 ACCTTATGATCATGCAAGAGAACTCAGACGGATTGTTAGGCACGAT 150
57 ASpglytyrValserThrserleuSerAlaHisleuIaI 73
151 GATGGATATGTTCCACCTCAATTAGTTGAGAGATGCCACCAAGGG 200
73 yGlnSerIleleuSerGlytyrSerThrtyrTyrlleTyValIleAlaI 90
201 TCAACATATGTTGCTGTCATCTACTATTATATATATATATATACCA 250
90 hAlaIProAsnMetPheAsnValAsnAspValIleuGlyValItyrSerPro 106
251 CTGACCCCAACATGTTTAACGTATATGATGTATTAGGGCATACAGCCCT 300
107 HisProtyrGlnGlnGlnValSerAlaLeuGlyGlyIleProtyrSerG 123
301 CATCAATGATGACAGAAAGTTTCTGCTTGTAGTGGGATTCATACCTCCA 350
123 nIleTyrglytyrTyrglyValAsnPhcglyValIleAspGluArgLeuH 140
351 AATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
140 IsAtGAsnAtGgLUtYrArGAspArGtYrTyrgAsnLeuAsnIleAla 156
401 ATCTAATATAGGGCTACAGAGATGATATTTACAGTAACTTATGATGCT 450
157 ProIaGluAspGlytyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 173
451 CCACGACAGAGTGGTATGATGATGATGATGATGATGATGATGATGATGAT 500
173 aTrpArGluGluProtyrIleHisIleAlaProGlnGlyCysGlyAsnS 190
501 TTGGAGGGAAGAGCCGATGATGATGATGATGATGATGATGATGATGATG 550
190 eRserArgthr 193
551 CTCACGATCA 561

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAF74431
seq_documentation_block:
ID AAF74431 standard; DNA; 347 BP.
XX
AC AAF74431;
XX
DT 08-MAY-2001 (first entry)
XX
DE Vibrio cholerae ctxA gene 347 bp region nucleotide sequence.
XX
KM Vibrio cholerae; ctxA gene; detection; identification; toxigenic; ds.
XX
OS Vibrio cholerae.
XX
PN ZA200001349-A.
XX
PD 27-DEC-2000.
XX
PF 16-MAR-2000; 2000ZA-0001349.
XX
PR 16-MAR-2000; 2000ZA-0001349.
XX
PA (UYPR-) UNIV PRETORIA.
XX
PI Theron J, Venter SN, Brozel VS, Du Preez M;
XX
DR WPI; 2001-191829/19.
XX
XX
XX Novel oligonucleotide primer useful for PCR amplification of toxigenic
XX PT Vibrio cholerae nucleic acid for detecting toxigenic Vibrio cholerae
XX PT organisms present in the sample
XX PS Example 2; Fig 7; 4pp; English.
XX
```

```
CC The present invention describes a method for amplifying (A1) toxigenic
CC Vibrio cholerae nucleic acid by: (a) performing a polymerase chain
CC reaction (PCR), by combining a medium containing toxigenic V. cholerae
CC nucleic acid with an oligonucleotide primer (1) which will bind to a
CC section of V. cholerae nucleic acid, in an environment suitable for
CC carrying out PCR; (b) combining the medium with a second primer which
CC will bind downstream to the first primer to a section of the
CC complementary strand of the V. cholerae nucleic acid; and (c) amplifying
CC corresponding sections of V.cholerae nucleic acid present in the medium.
CC The method is useful for detecting toxigenic V. cholerae organisms or
CC toxigenic V. cholerae nucleic acid, e.g., ctxA gene, in an environmental
CC water sample e.g., tap water sample, river water sample, lake water
CC sample or a sewage sample. The method is useful in PCR assays for
CC detecting a microorganism or nucleic acid of the microorganism in a
CC sample. (1) and amplification products are useful for detecting toxigenic
CC V. cholerae nucleic acid in a sample. The PCR method involving (1) is
CC rapid, economical, specific and sensitive for detecting toxigenic
CC V. cholerae in water samples. The present sequence represents a
CC region from the V. cholerae ctxA gene, which is used in an example from
CC the present invention.
SQ
Sequence 347 BP; 101 A; 66 C; 77 G; 103 T; 0 other;

alignment_scores:
Quality: 562.00 Length: 115
Ratio: 4.930 Gaps: 0
Percent Similarity: 99.130 Percent Identity: 91.304

alignment_block:
US-09-528-682-1 x AAF74431 ..
Align seg 1/1 to: AAF74431 from: 1 to: 347

7 ArgAlaAspSerArgProProAspGluIleIysArgSerGlyGlyLeu 23
1 CGGGCAGATTCTAGACCTCCTGATGAATAAAGCAGTCAGTGCTTAT 50
23 tProArgGlyHisAsnGlyThrPheAspArgGlyThrGlnMetAsnIleA 40
51 GCCAAGAGCAGAGTGAAGTACTTTGACCGAGTACTCAATGAAATATCA 100
40 snLeuTyrsPnHisAlaArgGlyThrGlnThrGlyPheValArgTyraSp 56
101 ACCTTATGATCATGCAAGAGAACTCAGACGGGATTTGTTAGCAGAT 150
57 ASpglytyrValserThrserleuSerleuArgSerAlaHisleuIaI 73
151 GATGGATATGTTTCCACCTCAATTAGTTGAGAGATGCCACCTTATG 200
73 yGlnSerIleleuSerGlytyrSerThrtyrTyrlleTyValIleAlaI 90
201 TCAACATATGTTGCTGTCATCTACTATTATATATATATATATACCA 250
90 hAlaIProAsnMetPheAsnValAsnAspValIleuGlyValItyrSerPro 106
251 CTGACCCCAACATGTTTAACGTATATGATGTATTAGGGCATACAGCTCT 300
107 HisProtyrGlnGlnGlnValSerAlaLeuGlyGlyIleProtyr 121
301 CATCAATGATGACAGAAAGTTTCTGCTTGTAGTGGGATTCATACCTCCA 345

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAT03448
seq_documentation_block:
ID AAT03448 standard; DNA; 777 BP.
XX
XX AAT03448;
XX
XX
XX 02-MAY-1996 (first entry)
XX
XX Shigella flexneri thermostable enterotoxin gene fragment Lf.
XX
XX
```


KW Thermolabile enterotoxin; subunit-A; subunit-B; detection; probe;
KM Ltf fragment; enterobacteria; ss.
XX Shigella flexneri.
XX R02031948-C1.
XX PD 27-MAR-1995.
XX PF 11-DEC-1991; 91RU-0016860.
XX PR 11-DEC-1991; 91SU-4016860.
PA (NIZH-) NIZHEGOROD EPIDEMIOLOGY MICROBIOLOG INST.
PI Masepa VN, Skoblo LE, Ulanova TI;
DR WPI; 1995-343066/44.
XX Ltf DNA fragment for recognition of thermolabile enterobacterium
PT enterotoxin gene - can be used as a probe for recognition of
PT Lf-toxin DNA gene
XX PS Claim 1; Columns 3-4; 5pp; Russian.
CC The present sequence is that of the Ltf fragment which encodes part
CC of the thermolabile enterotoxin subunit-A sequence and all of the
CC subunit-B sequence of Shigella flexneri. The DNA fragment is useful
CC as a probe for specific detection of toxigenic enterobacteria.
XX SQ Sequence 777 BP; 288 A; 141 C; 153 G; 195 T; 0 other;

alignment_scores:
Quality: 354.00 Length: 68
Ratio: 5.206 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 94.118

alignment_block:
US-09-528-682-1 x AAV03448 ..

Align seg 1/1 to: AAT03448 from: 1 to: 777

```
173 AATTPAAGLUGLUPROTPIIHHSHALAPROGLNGLYCYSGLYAS 189
|||||
3 GCCTGAGAGAAACCTGATTCATGCACCAAGGTTGGA 52
189 nuserATgThrlleThrgLYASpThrcYsAnGLuGLuThrGLAAsnL 206
|||||
53 TTCATCAAGACATTCACAGTGATCTGTATGAGAGGCCAGAAATC 102
206 euserThrlleThrgLYASpThrcYsAnGLuGLuThrGLAAsnL 222
|||||
103 TAGACACATATATCTCAGGAAATATCAATCAAAAGTTAAGAGCAGATA 152
223 PheserAspTYrGLInserGLuValAspIleTYrAsnArgILearGspGL 239
|||||
153 TTTTCGACTATCATCGACAGGTTGACATATATAACAGAAATCGGATGA 202
```

239 uLeu 240
||||
203 ATTA 206

seq_name: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAV62460

seq_documentation_block:
ID AAV62460 standard; DNA: 2385 BP.

XX AC AAV62460;
XX DT 15-FEB-1999 (first entry)
XX DE Helicobacter pylori ureb-Vibrio cholerae A2+B subunit gene fusion.

XX Vaccine; antigen: Ureb; antigen; toxin; diagnosis; gastritis;
KM ulcer; stomach cancer; ss.
XX OS Chimeric - Helicobacter pylori.
XX OS Chimeric - Vibrio cholerae.
XX Key Location/Qualifiers
FT CDS 1..1680
FT /*tag= a
FT /*product= ureb
FT CDS 1681..2256
FT /*tag= b
FT /*product= A2+B subunits
XX WO9844130-A1.
XX 08-OCT-1998.
XX 31-MAR-1998; 98WO-KR00073.
XX 31-MAR-1997; 97KR-0011951.
XX 31-MAR-1997; 97KR-0011950.
XX (DAEW-) DAEWOONG PHARM CO LTD.
XX Chol D, Jung H, Kim B, Park M, Shin S, Yu Y;
PI WPI; 1998-568279/48.
DR P-PSDB: AAW80599.
XX New chimeric proteins for use against Helicobacter pylori -
PT comprising an antigenic protein of H. pylori and A1 and B subunits
PT of Vibrio cholerae toxin, preferably produced by recombinant
XX techniques
XX Claim 3; Page 77-78; 102pp; English.
XX This is the nucleotide sequence of a fusion gene prepared by
CC ligating the ureb gene of Helicobacter pylori and the A2 and B
CC subunits genes of Vibrio cholerae toxin. It codes for a claimed
CC 749-amino acid fusion protein (see AAW80599). The invention relates
CC to: chimeric proteins comprising antigenic proteins of H. pylori
CC and A2 and B subunits of V. cholerae toxin; recombinant DNAs
CC encoding such chimeric proteins; recombinant expression vectors;
CC a process for preparing the chimeric proteins using transformed
CC microbial host cells; and to preventative and therapeutic vaccines
CC comprising the chimeric proteins for H. pylori-associated diseases
CC such as gastritis, gastric ulcer, duodenal ulcer and gastric cancer
CC (claimed). The chimeric proteins are designed to have excellent
CC immunogenicity, to be stable in the stomach, to penetrate the
CC mucous membrane of the intestine, and to stimulate production of
CC sIgA. They can additionally be used as active ingredients in
CC diagnostic kits for H. pylori infection, and for production of
XX anti-H. pylori antibody.
XX SQ Sequence 2385 BP; 779 A; 486 C; 509 G; 611 T; 0 other;

alignment_scores:
Quality: 252.00 Length: 112
Ratio: 3.111 Gaps: 3
Percent Similarity: 72.321 Percent Identity: 47.321

alignment_block:
US-09-528-682-1 x AAV62460 ..

Align seg 1/1 to: AAV62460 from: 1 to: 2385

```
137 GUARGLeuHISARgAsnARgGLuTYrARgAspARgTYrTYrARg..... 151
|||||
1557 CAGAAATATCATCTAAAGAAAGACATTCATCAACGACATACGCGTCAC 1606
```



```
152 .....AsnLeuAsnIleAlaProAlaGluAspGlyTyrA 163
      |||||
1607 TTGAAGTCAATTCTGAACCTA.....CCATGCTGTCGTGATGAC 1647
      |||||
163 rGleuAlaGlyPheProAlaAspHisGlnAla.TPrArgGluGluProTr 179
      ::::: ||| ::::: |||||
1648 AAAGAAGTAACTCTAAACCAACCAATAAAGTGAAGATTGGAAGAGCCGTG 1697
      ::::: ||| ::::: |||||
179 PLeHisHisAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrg 196
      ||||| ||||| ||||| |||||
1698 GATTCATCATGACACCGCGGTTGTGGGAATGCTCCAGATCATGATCA 1747
      ::::: ||||| ::::: |||||
196 lYAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArg 212
      ::::: ||||| ::::: |||||
1748 GTAATACTTCGATGAAAAAACCAAGCTAGGTGTAAATAATCTTGAC 1797
      ::::: ||||| ::::: |||||
213 GluTyrGlnSerLysValLysArgGlnIlePheSerAspTyrGlnSerG 229
      ||||| ||||| ||||| |||||
1798 GAATACCAATCTAAAGTTAAAGACAAATATTTCAGGCTATCAATCTGA 1847
      ::::: ||||| ::::: |||||
229 uValAspIleTyrAsnArgIleArgAspGluLeu 240
      ::::: ||||| ::::: |||||
1848 TATTGATACACATATAAGAAATTAGGATGATTA 1881
```


gb_est1:AI102743	-	79.00	143.45	96.21	594	AI102743	EST212032	Normalized	+	77.00	139.17	166.42	604	BF575414	602135645F1	NIH_MGC			
gb_est1:AM210136	+	79.00	142.25	98.62	604	AM210136	u151c10.y1	Rashpass	+	77.00	139.08	168.47	609	BI605467	RH71079.5	Sprime RH D			
gb_est2:BM285021	+	79.00	142.84	103.99	626	BM285021	kn93c10.y1	Ascaris	+	77.00	139.06	168.88	610	BI656330	RH63081.5	Sprime RH D			
gb_est1:BM571778	-	79.00	141.92	116.80	719	BM571778	R1ENK	full-16	+	77.00	138.76	175.49	626	AI125949	LP02919.5	Sprime LP D			
gb_est2:BG574777	+	79.00	141.24	127.70	719	BG574777	EST496155	GSTS	Solanum	+	77.00	138.32	185.56	650	BI260859	BOGC7719F	BOGC Bras		
gb_est2:BG353038	+	79.00	140.95	132.47	737	BG353038	602457485F1	NIH_MGC_16	+	77.00	137.97	194.09	670	AG014025	Homo sapiens	genom1			
gb_est1:AI664880	+	79.00	140.19	146.02	787	AI664880	wa86h08.x1	Soares	test	+	77.00	137.96	194.52	671	AM652315	Sheared DNA-6k2	TR		
gb_gss:AZ934810	+	79.00	140.16	146.57	789	AI2934810	BT_Ba000022	var B	Japl	+	77.00	137.22	213.73	715	AG652322	EST320109	tomato ra		
gb_est2:BE727170	+	79.00	139.93	151.00	805	BI2271210	602952296F1	NIH_MGC_6	+	77.00	136.35	239.02	771	BF305954	HVSMBB0012	12A06F Hord			
gb_est2:BE780666	+	79.00	135.79	257.91	1155	BE780666	601469344F1	NIH_MGC_6	+	77.00	135.94	252.00	799	BG418932	602559912F1	NIH_MGC			
gb_est2:BI855336	+	79.00	134.42	305.95	1296	BI855336	603382377F1	NCI_CGAP	+	77.00	134.77	292.76	884	BF280560	601811768F1	NIH_MGC			
gb_est1:AA129448	+	78.50	147.59	56.54	380	AA129448	zm85h03.s1	Stratagene	+	77.00	134.01	322.72	944	BE780560	601582575F1	NIH_MGC			
gb_est2:BG959006	+	78.50	145.46	76.33	457	TH1172	y242407.r1	Soares	fetal	+	77.00	133.23	356.77	1010	BG616292	602644507F2	NIH_MGC		
gb_est1:AI5558961	+	78.50	144.49	84.18	497	BG959006	MR4-rh1052-150201	-001	+	77.00	132.94	369.89	1035	AG063739	Pan troglodytes	DN			
gb_est2:BI155441	-	78.50	140.33	127.33	657	AI528991	u43e02.y1	Soares	mam	+	77.00	131.84	426.32	1139	BI522651	603195922F1	NIH_MG		
gb_est2:BG587737	-	78.50	140.07	143.45	712	BI147246	BI147246	unpublished	c	+	77.00	127.12	780.85	1713	gb_hic:AK0152261				
gb_est2:BI155441	-	78.50	140.07	148.26	728	BI155441	BI155441	unpublished	c	+	77.00	123.18	1.3e+03	2409	AK012494	Mus musculus	11 da		
gb_gss:BM533735	+	78.50	139.98	150.08	734	BG587737	EST489512	MHAM	Mediac	+	76.50	139.40	161.54	542	BM255705	K0728C06-3	NIA Mous		
gb_est2:BG914788	+	78.50	138.38	184.28	843	BI533735	BOHB2307F	BOHB	Brasid	+	76.50	137.96	194.36	614	BI605367	RH70956.5	Sprime RH D		
gb_est2:BG414440	-	78.50	137.97	194.09	875	BG490651	602520125F1	NIH_MGC_18	+	76.50	136.10	246.64	721	BG159770	de68d11.y1	Kirsche			
gb_est2:BI181562	+	78.50	137.06	218.30	945	BI181562	UNK-P-FN-a0-f-09-0-UNI		+	76.50	135.12	279.79	834	BM045368	603633364F1	NIH_MGC			
gb_est2:BG230664	+	78.50	136.83	224.84	964	BE295664	600944327F1	NIH_MGC_17	+	76.50	134.42	306.07	870	AL537110	AL537110	L1L1-FL013			
gb_est2:BF204688	+	78.50	127.88	708.33	2090	BF204688	601867526F1	NIH_MGC_1	+	76.50	130.47	489.07	1181	BG975390	602643041F1	NIH_MG			
gb_gss:AO909105	+	78.00	147.16	70.56	361	BB840688	BB840688	R1ENK	full-16	+	76.50	128.52	652.85	1390	AG097062	Pan troglodytes	DN		
gb_est1:AA803232	+	78.00	142.84	110.90	548	AA803232	GM043636	Sprime	GM Dros	+	76.50	127.03	790.20	1581	RC015622	Homo sapiens	Stm1		
gb_est1:AI1486065	+	78.00	141.86	117.87	571	AI1486065	EST244386	tomato	ovary	+	76.50	118.68	2.3e+03	3255	AK014599	Mus musculus	0 day		
gb_est2:BI306660	+	78.00	141.82	118.48	573	BI306660	NL_5_124	Drought	stres	+	76.50	118.32	2.4e+03	3357	AK005011	Mus musculus	adult		
gb_est1:AA802683	+	78.00	141.60	121.87	584	AA802683	GM06110	Sprime	GM Dros	+	76.00	139.67	156.14	485	AA662886	z1j6h04.s1	Soares	f	
gb_est2:BG655776	+	78.00	141.37	125.60	596	BG635776	SD13264	Sprime	SD Dros	+	76.00	138.22	187.14	550	BM277429	Tm-ad	40A06-SKPL	TJ R	
gb_est2:BI865854	+	78.00	141.35	125.91	597	BI838554	SD20866	Sprime	SD Dros	+	76.00	138.26	188.16	551	AO964439	RPCI-23-322E11	TV R		
gb_est2:BI206448	+	78.00	141.33	126.23	598	BI206448	EST524488	c10S	Lycopod	+	76.00	137.99	193.76	561	AO968780	RPCI-23-322P16	TJ R		
gb_est2:BG640826	+	78.00	141.31	126.54	599	BG640826	SD11981	Sprime	SD Dros	+	76.00	137.62	203.05	579	AQ038687	RC1111-138112	TV RP		
gb_est2:BI113506	+	78.00	141.19	128.42	605	BI113506	RE16438	Sprime	RE Dros	+	76.00	137.21	214.07	600	AW597203	EST318826	MHAM	Medi	
gb_est2:BI151326	+	78.00	141.14	129.37	608	BI151326	RE02704	Sprime	RE Dros	+	76.00	136.48	235.03	639	AG139018	Pan troglodytes	full		
gb_est1:AI1113376	+	78.00	140.87	133.81	622	AI1113376	GH08445	Sprime	GH Dros	+	76.00	136.12	246.02	659	BE729025	601583420F1	NIH_MGC		
gb_est2:BM318921	+	78.00	139.96	150.40	673	BM318921	K162a06.y1	Ascaris	sou	+	76.00	133.99	250.46	667	BG847824	1024019604.y1	C. re		
gb_est1:AI234208	+	78.00	139.89	151.73	677	AI234208	LP07529	Sprime	LP Dros	+	76.00	133.97	251.01	668	BB648070	BB648070	R1ENK	full	
gb_est1:AM037263	+	78.00	139.88	152.06	678	AM037263	EST274113	tomato	mixed	+	76.00	133.69	259.98	684	BB041966	BB041966	NIBH	Mochi	
gb_est2:BF479279	+	78.00	139.64	156.74	692	BF479279	I48-294279	Ice	Plant	+	76.00	133.59	263.37	690	BB648001	BB648001	R1ENK	full	
gb_est1:AL653906	+	78.00	139.41	161.46	701	AL653906	AL653906	XGC	gastrula	+	76.00	133.54	265.07	693	BI543385	SILA-C6	Sugar	Beet p	
gb_est2:BF308520	+	78.00	138.85	173.48	741	BF308520	601887667F1	NIH_MGC_17	+	76.00	133.49	266.77	696	AI598804	LS06d10.y1	NCI_CGAP			
gb_est2:BG972287	+	78.00	138.21	188.25	783	BG972287	602841418F1	NCI_CGAP	K	+	76.00	133.26	274.77	710	BM300036	MCR054D08	24600	Ice	
gb_gss:AG187693	+	78.00	137.51	205.99	832	BG847944	1024019H01.x1	C. reinh		+	76.00	133.23	275.92	712	AW475303	un6404.y1	Sugeno	m	
gb_est2:BE811552	+	77.50	143.78	93.20	443	BE811552	ISC006	G09F90629	ITRC	+	76.00	135.07	281.69	722	AW12834	JALSLC167	SHC	Litom	
gb_gss:AO653882	+	77.50	139.95	150.70	617	AO653882	Sheared DNA-9N16	TF	SH	+	76.00	134.91	287.49	732	BF458735	AT19609.5	Sprime	AT D	
gb_est1:AI833876	+	77.50	139.74	155.43	630	AI833876	605096F05.x2	G05	1	ND-T	+	76.00	134.74	293.92	743	AI478735	tm24601.x1	Soares	N
gb_gss:BM398288	-	77.50	136.74	175.97	685	BM398288	AG-ND-138P21	TF	1	ND-T	+	76.00	134.57	300.40	754	BM048736	603652338F1	NIH_MGC	
gb_est2:BF107273	-	77.50	138.55	160.18	686	BF107273	601824167F1	NIH_MGC_73	+	76.00	133.47	345.75	829	BI250960	6029933205F1	NCI_CGA			
gb_est1:AI200892	+	77.50	137.64	192.49	737	BI375192	601228346F1	NCI_CGAP	M	+	76.00	133.77	350.09	832	BI250850	BOGNO997F	BOGN	Birs	
gb_est2:BI6564800	+	77.50	137.64	202.44	753	AI200892	qf68c08.x1	Soares	test	+	76.00	133.75	379.03	836	BE311167	601957885F1	NCI_CGA		
gb_gss:CN501M17	+	77.50	136.08	247.44	862	BI656480	603288215F1	NCI_CGAP	M	+	76.00	131.92	421.85	948	BF025472	601670597F1	NIH_MGC		
gb_est2:BG9860735	+	77.50	132.39	397.16	1186	AI150348	Anopheles	gambie	GSS	+	76.00	130.50	506.20	1072	BM584844	AGENCORPT	6414078		
gb_gss:AO910494	+	77.00	143.57	94.71	413	AO910494	G5STC05730	Typanosoma		+	76.00	132.24	1.5e+03	2189	AK003538	Mus musculus	18 da		
gb_est1:AA033034	+	77.00	142.10	114.36	469	AA033034	m137a05.y1	Soares	mouse	+	75.50	138.52	159.25	450	BI095243	dac03a09.y1	NICHX	X	
gb_est2:BE455520	+	77.00	141.94	116.18	474	BE455520	WHE1287-1290	BIH	BOB25	+	75.50	138.16	189.51	506	BI095243	dac03a09.y1	NICHX	X	
gb_est2:BG711191	+	77.00	141.56	122.78	492	BG711191	602719609F1	NIB_MGC_60	+	75.50	137.78	199.03	523	BE018530	130542	MARC	IPRG	Su	
gb_gss:TM2960030	+	77.00	140.84	134.42	523	AI485332	T. brucei	sheared	genc	+	75.50	136.99	220.26	560	BM094774	sa121b07.y1	Gm-cl06		
gb_est2:BI333851	+	77.00	139.92	151.13	566	BI333851	GM25375	Sprime	GM Dros	+	75.50	136.99	220.26	560	BM094774	sa121b07.y1	Gm-cl06		
gb_est1:AI297329	+	77.00	139.52	159.12	586	AI297329	LP11607	Sprime	LP Dros	+	75.50	136.05	248.22	607	BE922042	UT-M-B21	bee-d	-03-0	
gb_est1:AI257564	+	77.00	139.46	160.33	589	AI257564	LP05968	Sprime	LP Dros	+	75.50	136.02	249.44	609	BE922042	UT-M-B21	bee-d	-03-0	
gb_est1:AI1238573	+	77.00	139.44	160.73	590	AI257564	GH14662	Sprime	GH Dros	+	75.50	133.92	252.48	614	BI066942	BU066942	NIBH	Mochi	
gb_est2:BG6560447	+	77.00	139.42	161.14	591	BI066447	SD11465	Sprime	SD Dros	+	75.50	134.49	307.67	695	BF487338	AT22185	Sprime	AT D	
gb_est2:BG656546	+	77.00	139.38	161.95	593	BI066546	SD14283	Sprime	SD Dros	+	75.50	133.89	323.47	732	BF244029	HVSMF0002M21F	Hord		
gb_est1:AI1388847	+	77.00	139.31	163.57	597	AI1388847	GH19879	Sprime	GH Dros	+	75.50	133.55	342.38	754	BF687186	963087E02.y2	C. rei		
gb_est1:AA212882	+	77.00	139.17	163.57	604	AA212882	mm78e09.r1	Soares	mouse	+	75.50	133.01	366.90	790	AL228351	Tetradodon	mgroviati		
gb_est2:BI230460	+	77.00	139.17	166.42	604	BI230460	RE14142	Sprime	RE Dros	+	75.50	133.81	376.58	804	BE536902	601064702F1	NIH_MGC	</	

gb_est2.BF955435	75.50	130.52	490.56	824	BF495453	AT004362	5prtime	AT	Dros	gb_est1.BE170863	75.00	120.74	540.73	880	BF104659	6020315329P1	NIH_MGC	NIH_MGC	NIH_MGC
gb_g98.CNS02058	75.50	130.22	524.29	961	BF455624	6025359595P1	NIH_MGC	58	NIH_MGC	gb_est2.BE170863	75.00	129.83	547.14	947	BF578934	6020955148P1	NIH_MGC	NIH_MGC	NIH_MGC
gb_est2.BE616936	75.50	129.50	575.35	1005	BF141482	Tetraodon	nigrovittid			gb_est2.BE170863	75.00	129.33	587.89	994	BF578934	6020955148P1	NIH_MGC	NIH_MGC	NIH_MGC
gb_est2.BE616936	75.50	129.50	575.35	1005	BF141482	Tetraodon	nigrovittid			gb_est2.BE170863	75.00	129.33	587.89	994	BF578934	6020955148P1	NIH_MGC	NIH_MGC	NIH_MGC
gb_est2.BF933223	75.50	127.92	704.87	1027	BF616366	601279176P1	NIH_MGC	4	NIH_MGC	gb_est2.BE170863	75.00	129.31	589.65	996	BF1272750	6015670791P1	NIH_MGC	NIH_MGC	NIH_MGC
gb_est2.BF933223	75.50	126.88	805.02	1342	BF973233	602241502P1	NIH_MGC	4	NIH_MGC	gb_est2.BE170863	75.00	129.16	601.09	1009	BF177434	6013514068P1	NIH_MGC	NIH_MGC	NIH_MGC
gb_est1.BE127265	75.50	124.60	1.1e+03	1635	BF038306	601458125P1	NIH_MGC	6	NIH_MGC	gb_est2.BE170863	75.00	127.91	705.43	1124	BF073826	601570894P1	NIH_MGC	NIH_MGC	NIH_MGC
gb_est2.BG410293	75.00	141.70	120.30	341	BE127683	DEPA1432	Rat	LMBG	6	gb_est2.BE170863	75.00	126.44	852.41	1277	BE130763	601570894P1	NIH_MGC	NIH_MGC	NIH_MGC
gb_est2.B1952721	75.00	139.63	156.97	408	BG1401295	947051A011.x1	947 - 2	W		gb_est2.BE170863	75.00	120.73	1.8e+03	2091	BF494661	AV294661	Riken	NIH_MGC	NIH_MGC
gb_est1.AA569222	75.00	138.62	178.53	445	AA569223	nm13107.s1	NCI	CGAP	L1	gb_est1.BE170863	74.50	139.90	151.67	365	BB080814	BB080814	NIH_MGC	NIH_MGC	NIH_MGC
gb_est2.BP072267	75.00	138.50	181.51	450	BF072267	227104	MARC	2BOV	Bos	gb_est1.BE170863	74.50	137.88	196.07	434	BF072267	227104	MARC	2BOV	Bos
gb_est2.BG889378	75.00	138.07	188.74	462	BF469221	UT-M-BH3-aue-c-02-0-UT				gb_est2.BE170863	74.50	137.15	215.80	463	BF469221	UT-M-BH3-aue-c-02-0-UT			
gb_est2.BG889378	75.00	137.35	210.32	467	BF469221	UT-M-BH3-aue-c-02-0-UT				gb_est2.BE170863	74.50	137.15	215.80	463	BF469221	UT-M-BH3-aue-c-02-0-UT			
gb_est2.BF846612	75.00	136.76	226.84	523	BF846612	EST516400	CGTD	Solanum		gb_est1.A1956529	74.50	134.68	296.03	573	BF846612	EST516400	CGTD	Solanum	
gb_est1.A1057716	75.00	136.63	230.71	529	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1801106	75.00	136.63	230.71	529	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BF846612	75.00	136.58	232.01	531	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est1.AV433446	75.00	136.26	241.79	546	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BM30803	75.00	136.13	249.72	552	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BF846612	75.00	136.06	250.38	558	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BF846612	75.00	135.99	250.38	558	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BM087870	75.00	135.80	256.38	569	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BF846612	75.00	135.74	258.39	571	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BE61657	75.00	135.72	259.06	572	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BE61657	75.00	135.72	259.06	572	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BG082308	75.00	135.66	261.08	575	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BG635883	75.00	135.60	263.10	578	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est1.A155891	75.00	135.50	266.48	583	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1141751	75.00	135.50	266.48	583	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.BG640752	75.00	135.48	269.88	588	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_g98.AQ0947740	75.00	135.38	270.56	589	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1568003	75.00	135.15	278.78	601	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1568003	75.00	135.15	278.78	601	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.09	280.84	604	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.09	280.84	604	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF846612	QVO-EN0058	-261000-441			gb_est1.A1956529	74.50	134.07	320.09	604	BF846612	QVO-EN0058	-261000-441		
gb_est2.B1565491	75.00	135.07	281.53	605	BF8466														

gb_hlc:AK009776	73.00	119.13	2.2e+03	1688	AK009776	Mus musculus adult ma	gb_est2:BF6748206	72.50	126.13	879.73	836	BZ678206	601801205P1	NIH_MGC	836
gb_hlc:AK009794	73.00	119.13	2.2e+03	1688	AK009794	Mus musculus adult ma	gb_gss:A6767143	72.50	126.13	879.39	839	AZ678143	ENTL1787R	Entomobd	839
gb_hlc:AK010215	73.00	119.10	1.2e+03	1693	AK010215	Mus musculus adult ma	gb_est2:BE9037133	72.50	126.12	887.18	844	BE903733	601677344E1	NIH_MGC	844
gb_hlc:AK009346	73.00	119.07	2.2e+03	1696	AK009346	Mus musculus adult ma	gb_gss:AZ133378	72.50	126.12	887.18	844	AZ133378	OSNBB0115B2E1	CUG3	844
gb_hlc:AK009154	73.00	119.06	2.2e+03	1698	AK010154	Mus musculus adult ma	gb_est2:BI6319633	72.50	126.12	932.75	873	BI631963	603311605P1	NCI_CGAP	873
gb_hlc:AK009871	73.00	119.05	2.2e+03	1699	AK009871	Mus musculus adult ma	gb_est2:BI5478218	72.50	125.72	934.34	874	BI547821	602133138Z1	NIH_MGC	874
gb_hlc:AK009619	73.00	119.03	2.2e+03	1703	AK009619	Mus musculus adult ma	gb_est2:BI5478218	72.50	125.72	934.34	874	BI547821	602133138Z1	NIH_MGC	874
gb_hlc:AK010157	73.00	118.87	2.3e+03	1727	AK010157	Mus musculus adult ma	gb_est2:BF3150811	72.50	125.56	953.43	885	BF315081	601902051F1	NIH_MGC	885
gb_est1:AMW07271	72.50	139.16	189.47	298	AMW07271	TGSESTZ86b07.y1.TGHR+	gb_est2:BE7834658	72.50	124.92	1.0e+03	937	BE783465	601471790E1	NIH_MGC	937
gb_est1:AM654546	72.50	137.85	197.06	306	AM654546	WTGZT7c01.x1.NCI_CGAP_P1	gb_est2:BF3134749	72.50	124.92	1.0e+03	937	BF313474	601901696P1	NIH_MGC	937
gb_est1:AM661177	72.50	135.44	268.52	377	AM661177	h180a07.x1.Sources.NFLI	gb_gss:CMS06V94	72.50	124.88	1.0e+03	940	AL146852	T3 end of clone AX	AX	940
gb_est1:AL6011188	72.50	135.20	277.01	385	AL6011188	DKRZP13300640.y1.313	gb_est2:BF1372725	72.50	124.88	1.0e+03	940	BF137272	6016508879P1	NIH_MGC	940
gb_est1:BE0296644	72.50	134.61	298.52	405	BE0296644	KP33a03.y1.TBN95TW-SF	gb_est1:BE216710	72.50	124.72	1.1e+03	953	BE216710	HV.CEB0011FO1F	Horz	953
gb_est1:BE0296644	72.50	134.56	299.81	406	AM848045	58930 MARC ABOV Bos ta	gb_est2:BF3315285	72.50	124.61	1.1e+03	956	BF315285	60189159P1	NIH_MGC	956
gb_est1:AA5891616	72.50	134.56	300.91	407	AA5891616	v163907.r1.Beddington	gb_gss:CMS06V94	72.50	124.61	1.1e+03	956	BF315285	T7 end of clone AX	AX	956
gb_est2:BF7156548	72.50	134.53	301.91	408	BF715628	ma09g1.y1.Sources.NME	gb_est2:BF3134737	72.50	124.56	1.1e+03	966	BF313477	601901012E1	NIH_MGC	966
gb_est2:BG6071104	72.50	134.00	322.99	427	BG607104	MHE244E9.D08.H15ES Trif	gb_est2:BF3136832	72.50	124.43	1.1e+03	977	BF313683	601903343E1	NIH_MGC	977
gb_est2:RS1118	72.50	133.60	339.95	442	RS1118	yg71a01.r1.Sources Infant	gb_est2:BF9750749	72.50	123.91	1.2e+03	1022	BF975049	602246533E1	NIH_MGC	1022
gb_est2:BE641132	72.50	133.02	366.51	465	BE641122	hy20h12.x1.NCI_CGAP_GC	gb_est2:BE9890571	72.50	123.91	1.2e+03	1040	BE989051	601446332P1	NIH_MGC	1040
gb_est1:AM081352	72.50	132.84	374.73	472	AM081352	xc41b11.x1.NCI_CGAP_CC	gb_est2:BF3636158	72.50	123.33	1.3e+03	1075	BF363158	602017618E1	NCI_CGAP	1075
gb_gss:AO8711102	72.50	132.53	390.13	485	AO8711102	nbe00042M3F CUG1 Rice	gb_est2:BF4043034	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est2:BF6281817	72.50	132.48	392.52	487	BF628187	HVSMEO00M3CF22F Hordeum	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_gss:AMQ490618	72.50	132.46	393.72	488	AMQ490618	RPCI1-71-270H14_TV_RPCS	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est2:BM428969	72.50	132.27	403.33	496	AMQ490618	RPCI1-71-270H14_TV_RPCS	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est1:AV9255043	72.50	132.15	409.37	501	AMQ490618	RPCI1-71-270H14_TV_RPCS	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est1:AI9775543	72.50	131.97	419.10	509	AMQ490618	RPCI1-71-270H14_TV_RPCS	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est1:BE3241115	72.50	131.97	419.10	509	BE3241115	140730 MARC IPig Sus s	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est2:BE3858488	72.50	131.96	421.55	511	BE3858487	7g37b01.x1.NCI_CGAP_Bt	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_gss:TA9A114107	72.50	131.43	447.49	532	AL1459572	T. brucei sheared genc	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est1:AMU149142	72.50	131.44	448.79	533	AMU149142	xf10h09.y1.NCI_CGAP_Ki	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est1:AMU153621	72.50	131.42	449.94	534	AMU153621	AMU153621.NT2NP3 Homo s	gb_est2:BM4105149	72.00	133.57	341.60	406	AL274342	vc01a12.r1.Sources	RI	406
gb_est1:AA3932020	72.50	131.16	465.06	536	AA3932020	LD101971.Sprline LD Dros	gb_gss:AO4999711	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI9948438	72.50	130.93	477.75	556	BI9948438	60112b08.y3.Campbell1 ze	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BE2622282	72.50	130.93	485.41	562	BE2622282	60112b08.y3.Campbell1 ze	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est1:AI1137350	72.50	130.78	487.98	564	AI1137350	UI-R-C2P-09-C-08-0-0U1	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI6156630	72.50	130.72	491.83	567	BI6136250	SD18361.Sprline SD Dros	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BF2858094	72.50	130.68	494.41	569	BF2858094	HVSMEO0014M07F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est1:AMU608008	72.50	130.38	513.86	584	AMU60808	da25b08.y1.Xenla 1LDic	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est1:AL120212	72.50	130.30	519.09	588	AL120212	DKRZP61F197.r1.761 (S	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI066944	72.50	130.26	521.71	590	BI066944	BI066944 NIBB Moch11 r	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI617647	72.50	130.13	530.91	597	BI617647	RM48046.Sprline RH Dros	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI986201	72.50	129.95	542.82	600	BI986201	3162-69 Mouse E14.5 re	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BM262967	72.50	129.95	542.82	606	BM262967	da101f12.y1.NICHD XGc	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BE807617	72.50	129.78	554.82	615	BE807617	sa28b11.y1.Gm-c1061 G1	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BG9987504	72.50	129.75	557.50	617	BG9987504	MR4-H11052-230401-004-	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est1:BM1615148	72.50	129.45	583.15	636	BM161504	BM161504 RIKEN full1-16	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI9594022	72.50	129.32	588.59	640	BI959422	RM1356.Sprline RH Dros	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est1:AV940653	72.50	128.93	618.85	667	AV940653	AV940653.K. Sato unput	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BG5411962	72.50	128.83	635.79	667	BG541962	602565465F1 NIH_MGC_77	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_gss:AG167128	72.50	128.81	638.57	669	AG167128	Pan troglodytes Dm4.	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI0451419	72.50	128.78	631.36	671	BI0451419	BI0451419 NIBB Moch11 r	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BE5719100	72.50	128.57	652.41	686	BE577314	L48-2140M3 rice Plant 1	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	128.22	653.67	701	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	128.14	665.10	709	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	128.09	669.11	712	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	128.06	692.28	714	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	127.85	709.61	727	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	127.85	709.61	727	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	127.85	709.61	727	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	127.85	709.61	727	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	127.85	709.61	727	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	127.85	709.61	727	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	127.85	709.61	727	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2_H03_77A	NIH_MGC	529
gb_est2:BI954900	72.50	127.85	709.61	727	BI954900	HVSMEO0020E21F Hordeum	gb_est2:BM3013307	72.00	130.51	505.77	529	AO499971	HS_-5211_A2		

gb_est2:BI355584	+	129.50	575.29	577	BI355584	GMJ1465.5ptime	GM Dros	gb_est2:BI375821	+	128.88	623.24	609	BI375821	RE63703.5ptime	RE D
gb_est2:BI566656	+	129.50	575.29	577	BI566656	RH36362.5ptime	RH Dros	gb_est2:BG653613	+	128.86	624.75	610	BG653613	SD13069.5ptime	SD D
gb_est2:BI566913	+	129.50	575.29	577	BI566913	RH36711.5ptime	RH Dros	gb_est2:BI353373	+	128.86	624.75	610	BI353373	GM33433.5ptime	GM D
gb_est2:BI590427	+	129.50	575.29	577	BI590427	RH0404.5ptime	RH Dros	gb_est2:BI604451	+	128.86	624.75	610	BI604451	RH69805.5ptime	RH D
gb_est2:BI623004	+	129.50	575.29	577	BI623004	RH04792.5ptime	RH Dros	gb_est2:BI628413	+	128.86	624.75	610	BI628413	RH56794.5ptime	RH D
gb_est2:BI623038	+	129.46	578.25	579	BI623038	RE08907.5ptime	RE Dros	gb_gss:AA1103967	-	128.86	626.27	611	AA1103967	RPCT-23-33618.TV RP	
gb_est2:BI571612	+	129.46	578.25	579	BI571612	RH06760.5ptime	RH Dros	gb_est2:BI4105892	+	128.82	627.79	612	BI4105892	GH46044.5ptime	GH D
gb_est2:BI607085	+	129.46	578.25	579	BI607085	RH74077.5ptime	RH Dros	gb_est2:BI616597	+	128.82	627.79	612	BI616597	GH46874.5ptime	GH D
gb_est2:BI613071	+	129.46	578.25	579	BI613071	RH50039.5ptime	RH Dros	gb_est2:BI613588	+	128.71	635.42	617	BI613588	GH48275.5ptime	GH D
gb_est2:BI628575	+	129.46	578.25	579	BI628575	RH66233.5ptime	RH Dros	gb_est2:BI1108913	+	128.71	636.94	618	BI1108913	GH48275.5ptime	GH D
gb_est2:BI630170	+	129.46	578.25	579	BI630170	RH5154.5ptime	RH Dros	gb_est2:BI636144	+	128.71	636.94	618	BI636144	SD17724.5ptime	SD D
gb_est2:BI486009	+	129.44	579.74	580	BI486009	RE63374.5ptime	RE Dros	gb_est2:BI1135881	+	128.69	638.47	619	BI1135881	GH13768.5ptime	GH D
gb_est2:BI565656	+	129.44	582.70	582	BI565656	RH6463.5ptime	RH Dros	gb_est2:BI1107605	+	128.63	643.07	622	BI1107605	GH05371.5ptime	GH D
gb_est2:BI617772	+	129.40	582.70	582	BI617772	RH48208.5ptime	RH Dros	gb_est2:BG640126	+	128.54	650.75	627	BG640126	SD11080.5ptime	SD D
gb_est2:BI1137866	+	129.38	584.19	583	BI1137866	GH10625.5ptime	GH Dros	gb_est2:BI587649	+	128.54	650.75	627	BI587649	RH28385.5ptime	RH D
gb_est2:BI1294636	+	129.38	584.19	583	BI1294636	LP08003.5ptime	LP Dros	gb_est2:BI637553	+	128.54	650.75	627	BI637553	SD19604.5ptime	SD D
gb_est2:BI367487	+	129.38	584.19	583	BI367487	RE3235.5ptime	RE Dros	gb_est2:BI604572	+	128.52	652.29	628	BI604572	RH69967.5ptime	RH D
gb_gss:A0652368	-	129.38	584.19	583	A0652368	Sheared DNA-22E20.TP		gb_est2:BI609509	+	128.52	652.29	628	BI609509	RH41304.5ptime	RH D
gb_est2:BI611934	-	129.36	585.16	584	BI611934	RH48419.5ptime	RH Dros	gb_est2:BI227916	+	128.39	663.10	635	BI227916	RE25307.5ptime	RE D
gb_est2:BI109001	-	129.34	587.16	585	BI109001	GH07487.5ptime	GH Dros	gb_est2:BI574395	+	128.38	664.65	636	BI574395	RH23406.5ptime	RH D
gb_est2:BI257321	+	129.34	587.16	585	BI257321	LP05644.5ptime	LP Dros	gb_est2:BI354513	+	128.34	667.75	638	BI354513	GM27103.5ptime	GM D
gb_est2:BI257913	+	129.34	587.16	585	BI257913	LP06415.5ptime	LP Dros	gb_est2:BI358147	+	128.34	667.75	638	BI358147	RE44832.5ptime	RE D
gb_est2:BI215760	+	129.34	587.16	585	BI215760	RE22362.5ptime	RE Dros	gb_est2:BI630897	+	128.34	667.75	638	BI630897	RH60228.5ptime	RH D
gb_est2:BI585930	+	129.34	587.16	585	BI585930	RH2635.5ptime	RH Dros	gb_est2:BI635454	+	128.32	669.30	639	BI635454	RH8725.5ptime	RH D
gb_est2:BI613170	+	129.34	587.16	585	BI613170	RH42267.5ptime	RH Dros	gb_est2:BI369651	+	128.30	670.86	641	BI369651	RE35902.5ptime	RE D
gb_est2:BI483130	+	129.32	588.65	586	BI483130	RE65971.5ptime	RE Dros	gb_est2:BI4369674	+	128.28	672.41	644	BI4369674	RE60533.5ptime	RE D
gb_est2:BI575792	+	129.30	590.14	587	BI575792	RH33095.5ptime	RH Dros	gb_est2:BI630847	+	128.23	677.08	645	BI630847	RH60173.5ptime	RH D
gb_est2:BI619132	+	129.30	590.14	587	BI619132	RH50109.5ptime	RH Dros	gb_est2:BI635654	+	128.21	678.64	645	BI635654	RE47160.5ptime	RE D
gb_est2:BI623454	+	129.28	591.63	588	BI623454	RE31296.5ptime	RE Dros	gb_est2:BI585421	+	128.21	678.64	645	BI585421	RE47987.5ptime	RH D
gb_est2:BI614087	+	129.28	591.63	588	BI614087	RH43379.5ptime	RH Dros	gb_est2:BG24452	+	128.20	680.20	646	BG24452	GM0400609F1.NIH.MGC	
gb_est2:BI6161073	+	129.26	593.13	589	BI6161073	RH50014.5ptime	RH Dros	gb_est2:BI674854	+	128.20	680.20	646	BI674854	RE22538.5ptime	RE D
gb_est2:BI620773	+	129.26	593.13	589	BI620773	RH51987.5ptime	RH Dros	gb_est2:BI564322	+	128.20	680.20	646	BI564322	RH37211.5ptime	RH D
gb_est2:BI569515	+	129.22	596.12	591	BI569515	RH11606.5ptime	RH Dros	gb_est2:BI581327	+	128.16	683.32	648	BI581327	RH63939.5ptime	RH D
gb_est2:BI593515	+	129.22	596.12	591	BI593515	RH11606.5ptime	RH Dros	gb_est2:BI62780	+	128.14	684.89	649	BI62780	RE20569.5ptime	RE D
gb_est2:BI623780	+	129.17	600.61	594	BI623780	RH54701.5ptime	RH Dros	gb_est2:BI214609	+	128.04	689.59	652	BI214609	RC3-BMD034-290200-0	
gb_est2:BI293014	+	129.17	600.61	594	BI293014	GH16076.5ptime	GH Dros	gb_est2:AMW93656	+	128.04	694.30	655	AMW93656	RC3-BMD034-290200-0	
gb_est2:BI582298	+	129.13	603.61	596	BI582298	RH20266.5ptime	RH Dros	gb_est2:BI626367	+	128.00	699.45	657	BI626367	AV916004.AV916004.K.Sato un	
gb_est2:BI590891	+	129.13	603.61	596	BI590891	RH04578.5ptime	RH Dros	gb_est2:AV916004	+	127.98	699.45	657	AV916004	AV916004.AV916004.K.Sato un	
gb_est2:AM219159	+	129.11	605.11	597	AM219159	EST101641.tomato root		gb_est2:BG819271	+	127.98	716.42	659	BG819271	GM027814429F1.NCI.CGA	
gb_est2:BI571601	+	129.11	605.11	597	BI571601	RH06807.5ptime	RH Dros	gb_est2:BI606366	+	127.95	702.18	660	BI606366	RH72739.5ptime	RH D
gb_est2:BI520148	+	129.09	606.62	598	BI520148	LD40208.5ptime	LD Dros	gb_est2:BI629069	+	127.89	706.92	663	BI629069	GM03615495F1.NIH.MGC	
gb_est2:AM190778	-	129.09	606.62	598	AM190778	zpb8904.sl StrataGene		gb_est2:BM005535	+	127.88	706.92	663	BM005535	GM03615495F1.NIH.MGC	
gb_gss:AA298187	-	129.07	608.12	599	AA298187	IM0335C21F.Mouse 10kb		gb_est2:BG708166	+	127.86	708.50	664	BG708166	GM02671640F1.NIH.MGC	
gb_est2:BI574540	+	129.07	608.12	599	BI574540	RH24469.5ptime	RH Dros	gb_gss:BI396352	+	127.86	710.08	664	BI396352	AG-ND-158B13.TF.ND-	
gb_est2:BI626576	+	129.05	609.63	600	BI626576	RH67559.5ptime	RH Dros	gb_est2:BI607153	+	127.79	716.42	659	BI607153	RH74157.5ptime	RH D
gb_est2:BI576256	+	129.05	609.63	600	BI576256	RH32806.5ptime	RH Dros	gb_est2:BG530416	+	127.62	732.36	670	BG530416	GM0259167F1.NIH.MGC	
gb_est2:BI609698	+	129.05	609.63	600	BI609698	RH14518.5ptime	RH Dros	gb_est2:BI669167	+	127.62	732.36	670	BI669167	GM0259167F1.NIH.MGC	
gb_gss:AA215477	-	129.05	609.63	600	AA215477	Sheared DNA-79C6.TR SH		gb_est2:BG428547	+	127.62	732.36	670	BG428547	GM0259167F1.NIH.MGC	
gb_est2:BI216490	+	129.03	611.13	601	BI216490	RE23636.5ptime	RE Dros	gb_gss:AC004489	+	127.53	740.37	664	AC004489	Pan troglodytes DNA	
gb_est2:BI566886	+	129.03	611.13	601	BI566886	RH36403.5ptime	RH Dros	gb_est2:BI551634	+	127.52	741.98	665	BI551634	GM0194289F1.NIH.MGC	
gb_est2:BI593230	+	129.03	611.13	601	BI593230	RH11204.5ptime	RH Dros	gb_gss:AA052213	+	127.37	756.48	694	AA052213	SHoared DNA-6K1.TR	
gb_est2:BI613438	+	129.03	611.13	601	BI613438	RH61012.5ptime	RH Dros	gb_gss:BH511041	+	127.35	758.10	700	BH511041	AV921119.AV921119.K.Sato un	
gb_est2:BI570349	+	129.01	612.64	602	BI570349	RH02657.5ptime	RH Dros	gb_est2:AV921119	+	127.22	771.08	703	AV921119	RPCT-24-399L18.TV R	
gb_est2:BI614035	+	128.99	614.15	603	BI614035	RH43317.5ptime	RH Dros	gb_gss:BH081005	+	127.22	771.08	703	BH081005	GM0279964F1.NIH.CGA	
gb_est2:BI616162	+	128.99	614.15	603	BI616162	RE10253.5ptime	RE Dros	gb_est2:BI646234	+	126.88	805.48	724	BI646234	EST399009.tomato br	
gb_est2:BI579702	+	128.99	614.15	603	BI579702	RE73772.5ptime	RE Dros	gb_est2:BG432480	+	126.88	805.48	724	BG432480	EST399009.tomato br	
gb_est2:BI619702	+	128.99	614.15	603	BI619702	RH50789.5ptime	RH Dros	gb_est2:BI1109262	+	126.29	825.36	736	BI1109262	GM02897238F1.NCI.CGA	
gb_est2:BI621437	+	128.99	614.15	603	BI621437	RH52779.5ptime	RH Dros	gb_est2:BM411046	+	126.13	885.92	772	BM411046	EST585373.tomato br	
gb_est2:BI587118	+	128.97	615.66	604	BI587118	RH27387.5ptime	RH Dros	gb_est2:AL515310	+	126.13	885.92	772	AL515310	LFI.NPFL006	
gb_est2:BI589960	+	128.97	615.66	604	BI589960	RH03488.5ptime	RH Dros	gb_est2:BI697087	+	126.13	885.92	772	BI697087	GM0348966F1.NCI.CGA	
gb_est2:BI621050	+	128.97	615.66	604	BI621050	RH53325.5ptime	RH Dros	gb_est2:BM111155	+	126.09	901.28	781	BM111155	EST585691.Potato ro	
gb_est2:BI623729	+	128.97	615.66	604	BI623729	RH65059.5ptime	RH Dros	gb_est2:BF051108	+	125.99	902.99	782	BF051108	EST3436283.tomato de	
gb_est2:BI626259	+	128.97	615.66	604	BI626259	RH65111.5ptime	RH Dros	gb_gss:CN501510	+	125.88	915.01	800	CN501510	RE400000.algrovial	
gb_est2:BI626259	+	128.95	617.18	605	BI626259	RE59948.5ptime	RE Dros	gb_est2:BI547167	+	125.76	933.99	811	BI547167	GM0190412F1.NIH.MGC	
gb_est2:BI626259	+	128.95	617.18	605	BI626259	RH21950.5ptime	RH Dros	gb_gss:AA126208	+	125.56	953.10	811	AA126208	OSNBNB0075F07.CUGI	
gb_est2:BI626259	+	128.95	617.18	605	BI626259	RH61545.5ptime	RH Dros	gb_est2:BI735368	+	125.09	1.0e+03	845	BI735368	GM0302863F1.NIH.MGC	
gb_est2:BI633812	+	128.95	617.18	605	BI633812	RH33051.5ptime	RH Dros	gb_gss:BH502834	+	125.02	1.0e+03	845	BH502834	BOOVA24TF.BOV.Bras	
gb_est2:BI633812	+	128.95	617.18	605	BI633812	RH33051.5ptime	RH Dros	gb_est2:BG829433	+	124.31	1.1e+03	888	BG829433	GM0276372F1.NCI.CGA	
gb_est2:BI605494	+	128.93	618.69	606	BI605494	RH71118.5ptime	RH Dros	gb_est2:BI142844	+	124.11	1.1e+03	920	BI142844	GM02988095F1.NCI.CGA	
gb_est2:BI607017	+	128.93	618.69	606	BI607017	RH73999.5ptime	RH Dros	gb_gss:CMS07MUP	-	124.09	1.2e+03	921	CMS07MUP	T3 end of clone	

gb_est2:BG034236	-	72.00	123.70	1.2e+03	953	1	BM034236	602302494F1 NIH_MGC_87
gb_est1:BE260590	-	72.00	123.44	1.3e+03	905	1	BE260590	601146069F1 NIH_MGC_19
gb_est1:AL560901	+	72.00	123.11	1.3e+03	1073	1	AL560901	AL560901 LTI_NFL010E
gb_est2:BF570467	+	72.00	123.02	1.3e+03	1011	1	BF570467	602186268F1 NIH_MGC_4
gb_gss:CN504453	+	72.00	122.46	1.4e+03	1061	1	AL290496	Tetradon nigroviridis
gb_gss:AL182715	+	72.00	121.66	1.4e+03	1137	1	AG182715	Pan troglodytes DNA
gb_est2:BM106426	-	72.00	121.41	1.6e+03	1162	1	BM106426	602961945F1 NCI CGAP
gb_gss:CN50616E	-	72.00	121.39	1.6e+03	1164	1	AL411384	T3 end of clone ANMA
gb_gss:AC110400	+	72.00	121.17	1.7e+03	1166	1	AG110400	Pan troglodytes DNA
gb_est2:BE250243	+	72.00	119.15	2.2e+03	1412	1	BE250243	600943234F1 NIH_MGC_1
gb_hic:AC002923	+	72.00	118.18	2.5e+03	1536	1	AK002923	Mus musculus adult ma
gb_est2:BE910236	+	71.50	134.63	3.0e+03	1747	1	BE910236	601503560F1 NIH_MGC_7
gb_est2:BF426161	+	71.50	134.13	3.17e+03	354	1	BF426161	sr94d06.y1 Gm-cl047 GI
gb_est1:BB809162	-	71.50	133.62	3.39e+03	370	1	BB809162	CM2-ANO076-090800-321
gb_est1:BB809162	-	71.50	133.28	3.54e+03	381	1	BB809162	BB809162 RIKEN full-l
gb_est1:BM107131	+	71.50	132.38	3.97e+03	412	1	BM107131	EST538472 P. infestans
gb_est1:AM011731	+	71.50	132.35	3.99e+03	413	1	AM011731	EST722553 Schistosoma
gb_est2:BM106426	-	71.50	131.85	4.25e+03	431	1	BM106426	510154 MARC 3BOV Bos t
gb_gss:AO459336	-	71.50	131.62	4.38e+03	440	1	AO459336	HS_5062.B2.H02.SP6E RH
gb_est2:BG096245	+	71.50	131.59	4.40e+03	441	1	BG096245	TALC1147G1DR TALC1 TR
gb_est2:BE602653	+	71.50	131.10	4.68e+03	460	1	BE602653	268030 MARC 3BOV Bos t
gb_est2:BE602653	+	71.50	131.05	4.71e+03	462	1	BE602653	OV2_23_B06.b1.A002 Ova
gb_gss:BM189981	-	71.50	130.66	4.95e+03	478	1	BM189981	ATX0A237F ATX0 Arabidop
gb_gss:BM189981	-	71.50	130.61	4.99e+03	480	1	BM189981	TJ1247F TAMU Arabidopst
gb_est2:BE356408	-	71.50	130.61	5.34e+03	503	1	BE356408	DGL_125_A04.b1.A002 Da
gb_est2:BG578943	+	71.50	130.05	5.36e+03	504	1	BG578943	dfo2c09.y1 Wellcome CH
gb_est1:AV921946	+	71.50	129.89	5.47e+03	511	1	AV921946	HS_AV921946 K. Scto unpub
gb_gss:AO822049	+	71.50	129.51	5.74e+03	528	1	AO822049	HS_5483.AL.D02.TTA RBC
gb_est2:BF076642	+	71.50	129.46	5.78e+03	530	1	BF076642	226249 MARC 2BOV Bos t
gb_est1:AL120787	-	71.50	129.25	5.94e+03	540	1	AL120787	unpubd07.x1 Stratagene
gb_est2:BE918075	-	71.50	129.18	5.95e+03	541	1	BE918075	CVL_1.F03.b1.A003 Ovar
gb_est2:BB355170	+	71.50	129.08	6.07e+03	548	1	BB355170	DGL_39_G08.b1.A002 Dar
gb_est2:BI444181	+	71.50	129.04	6.07e+03	548	1	BI444181	daa99h08.y3 Wellcome G
gb_est2:BF076642	+	71.50	129.04	6.10e+03	550	1	BF076642	225459 MARC 2BOV Bos t
em_estP1:BE355722	-	71.50	128.97	6.15e+03	553	1	BE355722	PIL_29.G08.b2.A002 Pat
gb_est1:AL183788	-	71.50	128.98	6.22e+03	557	1	AL183788	605095D09.x2 G05 - End
gb_gss:AO625722	-	71.50	128.79	6.30e+03	562	1	AO625722	CITRBI-EL-265415.FP CIT
gb_est2:BI444162	+	71.50	128.74	6.33e+03	564	1	BI444162	RC0-MT0003-140300-031
gb_est1:AV922060	+	71.50	128.66	6.40e+03	568	1	AV922060	AV922060 K. Scto unpub
gb_est2:BI446767	+	71.50	128.58	6.47e+03	572	1	BI468767	sa103b04.y1 Gm-cl050 G
gb_est2:BG677183	-	71.50	128.48	6.55e+03	575	1	BG677183	PMO-BT0587-110200-001
gb_est2:BI446767	-	71.50	128.48	6.55e+03	577	1	BI446767	daa7406.y1 NICHD XG
gb_est1:AL502082	+	71.50	128.30	6.70e+03	586	1	AL502082	AL503082 Hordeum vulg
gb_gss:BM450138	-	71.50	128.28	6.72e+03	587	1	BM450138	BOHEU037F BOHE Brassic
gb_est2:BI721504	+	71.50	128.09	6.98e+03	600	1	BI721504	1031056C12.y1 C. reinh
gb_est2:BI988031	+	71.50	128.03	6.98e+03	610	1	BI988031	3214-63 Mouse E14.5 re
gb_est2:BU021824	-	71.50	127.84	7.12e+03	610	1	BU021824	BU021824 ME01ISA CDNA
gb_gss:BM552517	-	71.50	127.71	7.24e+03	617	1	BM552517	BOHO039F BOHP Brassic
gb_est2:BI552815	+	71.50	127.56	7.33e+03	625	1	BI552815	603193791F NIH_MGC_95
gb_est2:BE612259	+	71.50	127.54	7.39e+03	627	1	BE612259	sr05h09.y1 Gm-cl049 GI
gb_gss:LB6766Y	+	71.50	127.52	7.41e+03	627	1	AL544540	Leishmania major Fried
gb_est1:AV914413	-	71.50	127.45	7.48e+03	631	1	AV914413	AV914413 K. Scto unpub
gb_est2:BU211081	+	71.50	127.37	7.55e+03	635	1	BU211081	EST529121 CTOS Lycoper
gb_est1:AL510715	+	71.50	127.30	7.62e+03	639	1	AL510715	Hordeum vulg
gb_est2:BU116291	+	71.50	127.19	7.73e+03	645	1	BU116291	unpublished
gb_est1:AL1668149	-	71.50	127.12	7.80e+03	649	1	AL1668149	605015E01.y1 G05 - End
gb_gss:AO657846	+	71.50	127.12	7.80e+03	651	1	AO657846	LERA077F LERA Arabidop
gb_est1:AV919449	-	71.50	127.09	7.84e+03	651	1	AV919449	AV919449 K. Scto unpub
gb_est1:AM576725	-	71.50	127.07	7.85e+03	652	1	AM576725	RC3-BT0316-070100-012
gb_est1:AV929342	+	71.50	127.02	7.91e+03	655	1	AV929342	AV929342 K. Scto unpub
gb_est1:AV937189	+	71.50	127.00	7.93e+03	656	1	AV937189	AV937189 K. Scto unpub
gb_est1:AV928354	+	71.50	126.98	7.94e+03	659	1	AV928354	AV928354 K. Scto unpub
gb_est1:AV920170	+	71.50	126.94	7.98e+03	662	1	AV920170	AV920170 K. Scto unpub
gb_est1:AV919180	+	71.50	126.81	8.03e+03	667	1	AV919180	K. Scto unpub
gb_est1:AV922466	+	71.50	126.81	8.12e+03	667	1	AV922466	AV922466 K. Scto unpub
gb_est2:BI726509	+	71.50	126.67	8.27e+03	675	1	BI726509	1031085F1.y1 C. reinh
gb_gss:AZ557122	+	71.50	126.67	8.27e+03	675	1	AZ557122	2M0223F24R Mouse 10kb
gb_est1:AV922508	+	71.50	126.53	8.42e+03	683	1	AV922508	K. Scto unpub
gb_est1:AL508601	-	71.50	126.25	8.72e+03	700	1	AL508601	AL508601 Hordeum vulg
gb_gss:AZ853830	-	71.50	126.15	8.84e+03	706	1	AZ853830	2M0157008F Mouse 10kb
gb_est2:BM442159	-	71.50	126.05	8.95e+03	712	1	BM442159	EBAN01.SQ002.G03_R
gb_est1:AL555533	-	71.50	125.92	9.12e+03	720	1	AL555533	GC6-gastru
gb_gss:BM104330	+	71.50	125.87	9.21e+03	726	1	BM104330	BOGEP47F BOGE Bras
gb_est2:BI934986	+	71.50	125.67	9.40e+03	736	1	BI934986	EST354753 tomato f1
gb_est2:BI761041	-	71.50	125.57	9.52e+03	742	1	BI761041	603048581F1 NIH_MGC
gb_est2:BG162884	-	71.50	125.54	9.95e+03	744	1	BG162884	d34.fell.y1 Wellcome
gb_gss:AZ506171	+	71.50	125.40	9.97e+03	753	1	AZ506171	IM03474704F Mouse 10
gb_gss:AZ506171	+	71.50	125.31	9.84e+03	759	1	AZ313185	OTJB0108P19F CUGI
gb_gss:AZ506171	+	71.50	125.14	1.0e+03	770	1	BM776200	602663158F1 NIH_MGC
gb_est2:BF105460	+	71.50	125.06	1.0e+03	776	1	BF105460	601822316F1 NIH_MGC
gb_est2:BF025801	+	71.50	124.95	1.0e+03	783	1	BF025801	601670222F1 NIH_MGC
gb_est2:BE618790	+	71.50	124.91	1.0e+03	786	1	BE618790	601462842F1 NIH_MGC
gb_est2:BE618790	+	71.50	124.72	1.1e+03	799	1	BE618790	602484487F1 NIH_MGC
gb_est2:BM110188	+	71.50	124.59	1.1e+03	808	1	BM110188	EST557724 potato ro
gb_est2:BM107131	+	71.50	124.42	1.1e+03	820	1	BM107131	602663158F1 NIH_MGC
gb_est2:BF078981	+	71.50	124.22	1.1e+03	834	1	BF078981	602663158F1 NIH_MGC
gb_gss:CN60202CF	+	71.50	124.15	1.1e+03	839	1	AL109368	Tetradon nigrovir
gb_gss:CN60202CF	+	71.50	124.15	1.1e+03	841	1	BF774600	GA_EB002110F1 Goss
gb_est2:BE6274600	+	71.50	124.06	1.2e+03	846	1	BE627460	602752512F1 NIH_MGC
gb_est2:BE6274600	+	71.50	124.06	1.2e+03	846	1	BE627460	601328631F1 NCI CGA
gb_est2:BE570511	+	71.50	124.02	1.2e+03	849	1	BE570511	HVSME0006014F Horz
gb_est2:BI957013	+	71.50	123.85	1.2e+03	861	1	BI957013	60348127F1 NIH_MGC
gb_est2:BI959808	+	71.50	123.84	1.2e+03	862	1	BF344603	602015224F1 NCI CGA
gb_gss:BM159743	-	71.50	123.69	1.2e+03	873	1	BM159743	ENTR595FTR Entamoeba
gb_gss:BM159743	-	71.50	123.69	1.2e+03	873	1	AZ547888	ENTR595FTR Entamoeba
gb_est2:BF682917	+	71.50	123.37	1.3e+03	898	1	BF682917	602117535F1 Soares
gb_est1:AL557984	+	71.50	123.26	1.3e+03	906	1	AL557984	AL557984 LTI_NFL008
gb_est1:AL530605	+	71.50	123.05	1.3e+03	923	1	AL530605	AL530605 LTI_NFL003
gb_est1:AL530605	+	71.50	122.99	1.3e+03	928	1	AM667230	GA_EA0007015 Gossy
gb_gss:CN60505MD	+	71.50	122.88	1.3e+03	934	1	AL458602	Tetradon nigrovir
gb_est2:BI95347	+	71.50	122.88	1.3e+03	937	1	BI95347	602044161F1 NIH_MGC
gb_gss:BM168164	+	71.50	122.53	1.4e+03	965	1	BM168164	602044161F1 NIH_MGC
gb_gss:BM168164	+	71.50	122.39	1.4e+03	969	1	BM168164	602044161F1 NIH_MGC
gb_est2:BE525474	+	71.50	122.39	1.4e+03	977	1	AL523474	LTI_NFL003
gb_est2:BE525474	+	71.50	122.39	1.4e+03	996	1	BE525474	602379415F1 NIH_MGC
gb_est2:BE525474	+	71.50	122.07	1.5e+03	1005	1	BE525474	601334727F1 NIH_MGC
gb_est2:BE525474	+	71.50	122.07	1.5e+03	1022	1	BE525474	601334727F1 NIH_MGC
gb_est2:BM444947	+	71.50	121.66	1.6e+03	1041	1	BM444947	AGNC00076387638
gb_hic:AF161316	+	71.50	121.61	1.6e+03	1045	1	AK004405	Mus musculus 18 da
gb_gss:CN504453	+	71.50	121.24	1.7e+03	1079	1	AL311426	Tetradon nigrovir
gb_est2:BE421046	+	71.50	120.72	1.8e+03	1129	1	BE421046	HMM005.C03 IREC HW
gb_est2:BE421046	+	71.50	120.52	1.8e+03	1149	1	BE421046	602296548F1 NIH_MGC
gb_hic:BM019754	+	71.50	119.03	2.2e+03	1307	1	AK009754	Mus musculus adult
gb_est2:BE512511	+	71.50	113.59	4.4e+03	2091	1	AK161516	Homo sapiens HSPC1
gb_est2:BE512511	+	71.50	113.59	4.4e+03	2091	1	BI027953	IL5-MT025F1-150201-3
gb_est2:BE512511	+	71.50	135.05	1.92e+03	231	1	BM35480	602653167F1 NIH_MGC
gb_est2:BE512511	+	71.50	135.05	1.92e+03	266	1	BM35480	602653167F1 NIH_MGC
gb_est2:BE512511	+	71.50	135.05	1.92e+03	266	1	BM35480	602653167F1 NIH_MGC
gb_est2:BE512511	+	71.50	1					

/clone_lib="NIH_MGC_89"
 /tissue_type="hypermphroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 246 a 231 c 229 g 233 t
 ORIGIN

alignment_scores:

Quality: 93.00 Length: 261
 Ratio: 0.838 Gaps: 13
 Percent Similarity: 42.529 Percent Identity: 19.540

alignment_block:

US-09-528-682-1 x BG171693 ..

Align seg 1/1 to: BG171693 from: 1 to: 939

```

4  AAGLEUTYATGALAAASPSERATGPRODASPGIUILEYSATGSEGT 20
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
227 AAGTGCATGCTGCTGAGATATTCCTGAGAAACCCCTGAAATGCTGA 276
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20  YGLYLEUMETPROARGLYHIS.....ASGLUYTYRPHASPARGLYT 35
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
277 TCCAAAATATGACAAAGACCCTTTGGAAAGTATTACTCCAGGCCAA 326
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
35  hrc1n.....Metasnlleasnlleutyasphlsalargly 47
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
327 AGAAGCACCAGAGCCATGGAACCTT.....GATGCCCTTAAGGA 367
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
48  ThGlnThGlypheValargTyraSPASPGLYTYValSerThrSerle 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
368 ACT.....GGATATATCAAGACTGAGT 390
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
64  useRleuATGserAlaHisleuNlaGlyInserIleleuSerGlyTyrS 81
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
391 G..... 391
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
81  erThrTyTyTyTyTyValIleAlaThAlaProAsnMetPheAsnVal 97
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
392 .....ATTCTGTGCTGAGTTCACCCCTTCTAGACTTCAGACC 430
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
98  AsnAspValleuGlyValTySerProHisProTyGlnGlnValSe 114
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 ACAGACACCTGCTGCCCATGTCCTCTGAGAGCTTGACGAGTGTCTCG 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
114  fAlaLeuGlyGlyIleProTySerGlnIleTyGly..... 126
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 GATAGTGGGCTGTGTAATTCGACAGTATGATGACAGCTATAGACA 530
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
127 .....TTPTYATGValAsnPhely..... 133
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
531 TGAATTTTTCATCTTCTGCGGACAGTTTCTTCTCATCTGTGAT 580
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
133 ..... 133
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
581 CCCTCTGCTACTCTGCTTCTACATCCTGTGTTCTAGGAAATGAAA 630
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 .....ValIleAspGluArgLeuHis 141
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
631 CGAAAGCGACAAATTCGCTGCAACCTGTGATGACAGAGATTTTTC 680
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
141  rGAsnArgGlu.....TYATGAspArgTyTyTyArgAsnLeuAsn 154
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
681 CTAACTCAGAAACATCAAGTTACTGTGAAAGGATCATGATCTTACGA 730
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
155  IleAlaProAlaGluAspGlyTyTyArgLeuAlaGlyPheProProAsp 171
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
731  GTA.....ACTGACAGGCTCTCGCAAGAGGGGGTTCCCGGAAACA 774

```

```

171  s.....GlnAlaTrpArgGluGluProTyr..... 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
775  TCGGATTTCCCAAGTTTGAGAGATGAGTCTGGGAAGAAATTCCTCT 824
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
180  .....IleHisAlaProGlnGlyCysGlyAsnSerSerArgThr 193
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
825  CCAGTTTGCTAGTTCGCCCTCAGGG...GGCAACACGCTTGACAG 871
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
194  IleThr...GlyAspPhrcysAsnGluGluThr 203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
872  CATACCCCTGGGACAAATTCCTGTAGAGACC 904

```

seq_name: gb_gss:A2698180

seq_documentation_block:

LOCUS A2698180 714 bp DNA linear GSS 24-JAN-2001
 DEFINITION RPCI-23-223D11.TU RPCI-23 Mus musculus genomic clone RPCI-23-223D11
 , DNA sequence.

ACCESSION A2698180
 VERSION A2698180.1 GI:12415007

KEYWORDS
 GSS.

SOURCE
 house mouse.

ORGANISM
 Mus musculus

REFERENCE
 AUTHORS

TITLE
 JOURNAL

COMMENT

Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 other GSSs: RPCI-23-223D11.TU
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhaoc@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC
 Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
 page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 223 row: D column: 11
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers

FEATURES
 source

1. 714
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-223D11"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site: 1:
 EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI methylase. Size
 selected DNA was cloned into the pBAC3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."
 BASE COUNT 191 a 179 c 181 g 163 t

ORIGIN

alignment_scores:

Quality: 90.50 Length: 80
 Ratio: 1.708 Gaps: 6
 Percent Similarity: 66.250 Percent Identity: 37.500


```

was primed with a Not I - oIioo(GT) primer [5],
TCGTACAGATCTCGAGTGGGGAGCGCCGCTTTTATTTTTTTTTTTTT
3') Double-stranded cDNA was ligated to Eco RI adaptors
[CATGATTCGGTACC], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pTZ19 vector.
Library constructed by Bob Barstead."

BASE COUNT      128 a      148 c      140 g      124 t      3 others
ORIGIN

alignment_scores:
    Quality:      89.00          Length:      117
    Ratio:        1.618         Caps:       8
    Percent Similarity: 47.009     Percent Identity: 29.060

alignment_block:
US-09-528-682-1 x AA473437 ..

Align seg 1/1   to: AA473437   from: 1   to: 543

125 TYRGIYTPTYrArGVAlasnpheglyVal.....lleasp...gluar 138
||| ||||| :||||| :||||| ||||| |||||
147 TATGAGTGTGTACATCTACGAGTATGSAGGCCCTCACAATCGATGGSGAGA 196
138 gleuhlsarGAsnArgGlutTyTarGsparGtyr..... 149
: :||| :|||
197 ATACATTCCTTTTAAGCAAGTATGCAGGCAAAATATATCCCTCCTGTCACMG 246
150 ..... TyTarGsAsnleuhsnllea 156
247 TAGCCAGACTACTGAGGCTGCACAGACCAATACCTTGAACTGAATGCACTA 296
157 PROAlAGL.....AspGLyTyTarGLEuAlAGlPheProProAs 170
||| ||| ||| :||| ||||| |||
297 CCAAGAGAACCTTGGGGCAATGTGGCTTGTCATCTGGGCTT..CCTTC 343
170 pHsglnAlatPrArygGluglu.....Prot 179
:||| ||||| :|||
344 CAACCAAAATTTGGAAACAGAGAGCAGCGAGAACCTCCGATACTCCCCA 393
179 rPLIHslHslALaprogIngly.....CySglaYnsEseraTy 192
394 GTCTCAGATGTCTCCGACAGGTTGGGCTTTGTGTATATTTCACAGTC 443
193 ThrIleThrgLYAspThrCYsaNgIngluThrGlnAsnLeuSerThrII 209
||||| : :||| ||||| |||||
444 CTtGGAAMAAGAGATGTGANCGGGGAGAAAGACAGCAAT..TCrNCACT 490

209 e 209
:
491 T 491

seq_name: gb_est2:B1910238

seq_documentation_block:
LOCUS B1910238 1147 bp mRNA linear EST 16-OCT-2001
DEFINITION M3036945FE1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5218236 5',
mRNA sequence.
ACCESSION B1910238
VERSION B1910238.1 GI:16173620
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1147)
AUTHORS Nih-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: rgabbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            CDNA Library Preparation: Life Technologies, Inc.
```


/lab_host="DHI0B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
 Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 309 a 379 c 404 g 333 t others

alignment_scores:
 Quality: 87.00 Length: 306
 Ratio: 0.690 Gaps: 16
 Percent Similarity: 41.176 Percent Identity: 20.261

alignment_block:
 US-09-528-682-1 x B1663851/rev ..

Align seg 1/1 to reverse of: B1663851 from: 1 to: 1426

```

2  G1YASpArgLeuTyArgAlaAspSer..... 10
876 GGGTCCAAACCTGTGAAGGAGCCGCCCTTTCGTGATCGCTAAACT 827
11 .....ArgProAspGluLeuAspArgSerGlyLeuMetP 24
826 GTTTCGCTTTCGGCCCCCT.....CCCCGGGAGACCTCTTCTCC 786
24  roArgGlyHisasnGluTyArgPheAspArgGlyThrGlnMetasnLeasn 40
785 CAAGGGGTGAT..... 775
41  LeuTyAspHisAlaArgGlyThrGlnThrGlyPheValArgTyArg.. 56
774 .....CACACTTCAGGGGGGGCTCTGTTCACGGCTTAACAGG 734
57 .....AspGlyT 59
733 GGGACAGGGGTCCCACTTCGCACACAGTCGTTGCCCAAGATGATGTC 684
59  yValSerThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyInSer 75
683 AATGTGCAATCATCTCTACGTCGTGTCTCTAGAACATCTCCC 634
76  IleLeu.....Se 78
633 TTGTTTCCTCGGGGACAAAGGAGGCTTTTCCCATGATGGGCCCA 584
78  rGlyTySerThrTyArgTyrIleTyValIleAlaThrAlaProasn.... 93
583 TGGTTCCTCTGTCGTTACGGGGGATAGTGTGATGGCCCTCGGCCAA 534
93 ..... 93
533 GACCATTAACGATGCCAAATTCAGATGTGATCATGTACTAGACT 484
94 .....MetPheasnValasnAspValLeuGly ValTySerP 106
483 TGCAGCGATGTTTCATTCACGTTCCGACTCCATGGGCTATGTTTGTG 434
106 roHisProTyArgLugln.....GluValSerAlaLeu... 116
433 CTCACAAAGACAGAGAACCCCATGTGTAATATGATCTCGCACCATCAT 384
117 .....GlyGlyLeuProTySerGlnIleTyGly.....TrpTyArg 129
383 CGAGACCTCGGGCTGAGATATACTCTTCTTACGTCACCAACAGGGAAG 334
129  g...ValasnPheGlyValIleAspGluArgLeuHisArgAsnArgGluT 145
333 TCCCGCTCGCTTTCATGCTAGTACACGCC..... 303
145  yArgAspArgTyArgAsnLeuAsnIleAlaProAlaGluAspGly 161

```

```

302 .....CAATCGCTAATCTAGCTCTTCGACAGACAGTC 270
162 TyArgLeuAlaGlyPheProAspHisGlnAlaTrpArgGluGluP 178
269 AACAGATTCAGGCTCTCAGC.....CGACGCTCTCC 238
178 otrPLeHis.....HisAlaProGlnGlyCys.....G 188
237 GATTCTCCACCCCTCCATTCAGACATGACGGGACACATGCCACAGTC 188
188 1YasnSerSerArgThrIleThrGlyAspThrCysasnGluThrGln 204
187 GCCTGAGAGGAGAACCTCGATGATGATGTGTCTGCGCTGCGGT 138
205 AsnLeuSerThrIleTyArgGluTyArgGlnSerIlyValysArgG 221
137 GCCCTCGCCCTGACTGCACCTGTCTGAGC...AGAAGTCAGCAGAGAG 91
221 nIlePheSerAspTy 226
90 TGTAGACGAGACTAC 75
seq_name: gb_est1:AA145382

```

seq_documentation_block:
 LOCUS AA145382 497 bp mRNA linear EST 11-FEB-1997
 DEFINITION m78g11.r1 Strata gene mouse heart (#937316) Mus musculus cDNA clone
 IMAGE:603620 5', mRNA sequence.
 ACCESSION AA145382
 VERSION AA145382.1 GI:1714756
 KEYWORDS EST
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 497)
 MARRA, M., HILLIER, L., ALLEN, M., BOWLES, M., DIETRICH, N., DUBUQUE, T.,
 SCHEIDENBERG, K., STEPCOE, M., TAN, F., UNDERWOOD, K., MOORE, B.,
 THELSTING, B., WYLIE, T., LENNON, G., SOARES, B., WILSON, R. and
 WATERSTON, R.
 The WashU-HHMT Mouse EST Project
 Unpublished (1996)
 CONTACT: Marra M/Mouse EST Project
 WASHU-HHMT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LINL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:369052
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 415.

FEATURES

source

```

1. 497
/organism="Mus musculus"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:603620"
/clone_lib="Strata gene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: heart; Vector: pBluescript SK-; Site.1:  

EcorI; Site.2: XhoI; Cloned unidirectionally. Primer:  

Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts.  

Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'  

adaptor sequence: 5' CTCGAGCTTTTCTTTTCTTTTCTTTT 3'."
BASE COUNT 158 a 118 c 126 g 95 t

```


ORIGIN

alignment_scores:

Quality: 85.50 Length: 180
Ratio: 0.940 Gaps: 8
Percent Similarity: 50.556 Percent Identity: 24.444

alignment_block:

US-09-528-682-1 x AA145382 ..

Align seg 1/1 to: AA145382 from: 1 to: 497

```

4 ArgLeuTyr...ArgAlaAspSerArgProProAspGluIleuLysArgse 19
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 AGAGTATACATACAGGACGAGCATATTCCTGCAGAGATATACATACAGCA 50
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
19 rgllyglLeuMetProArgglYhIsasnglutyRheasparglyThrg 36
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 GGCAGTACTCTCGCAGAGTATACATACA.....AGCAGGC 85
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
36 lmetasnIleasnLeuTyrAspHisAlaArgglyThrgInThrglyPhe 52
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
86 AGTACTCTCTCGCAGAGTATACATACAGGACGAGTACTCTCGCAGAGTAT 135
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
53 ValArgTyrAspAspGlyTyrValSerThSerLeuSerLeuArgSerAl 69
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
136 ACA.....TACAAAGTGGCAGTACTCTCGCAGAGTAT 167
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 ahIsleuAlaGlyGln.....SerIleLeuSerGlyTyrSert 82
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 ACATACAGGACGAGTACTCGCAGAGTATACATACAGGACGAGCA 217
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
82 hTyrTyrTyrIleTyrValIleAlaThraLProAsnmetPheasnValasn 98
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
218 CAGAGTATACATACAGGACGAGGAGTGTCTCTCGAGATTATACATACAGGC 267
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 AspValLeuGlyValTyr.....SerProHisPr 108
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
268 AGGCAGTATACAGTATATACATACAGGACGAGTGTCTCTCGAGATTATAC 317
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 oTyrgLugIngluVal.SerAla.....LeuglyglyIle 119
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 ATACAGGACGAGTACTCTCTCGCAGATATACATACAGGACGAGTACT 367
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 ProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGlyValIleas 136
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
368 CCTGCAGAGTATACATACAGAGTGG.....GCAGTGTCTCCA 402
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
136 rgluArgLeuIleArgAsnArgGluTyrTyrArgAspArgTyrTyrArgAsnL 153
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
403 GCGAGATATCATACAGGACGAGCATATCTCTCGCAGAGTATACATACAGG 452
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
153 euAsnIleAlaProAlaGluAspGlyTyrArgLeuAla 165
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
453 CAGGACGAGTACTCTCGCAGATATACATACAGGACGCA 490
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

seq_name: gp_est2:BG859250

seq_documentation_block: 686 bp mRNA linear EST 29-MAY-2001

LOCUS BG859250

DEFINITION 102406GC07.y1 C. reinhardtii CC-1690, normalized, lambda Zap II

ACCESSION BG859250

VERSION BG859250.1 GI:14240434

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii.

REFERENCE 1 (bases 1 to 686)

AUTHORS Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P., McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.

TITLE

Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants: project phase 2

JOURNAL

Unpublished (2000)

Contact: Charles Hauser

DCMB Box 91000

Duke University

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chauser@duke.edu.

Location/Qualifiers

1..686

/organism="Chlamydomonas reinhardtii"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone_1lb="C. reinhardtii CC-1690, normalized, Lambda Zap II"

/note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2: XhoI. This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. Bluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT

117 a 205 c 253 g 111 t

ORIGIN

alignment_scores:
Quality: 85.50 Length: 132
Ratio: 1.315 Gaps: 8
Percent Similarity: 49.242 Percent Identity: 29.545

alignment_block:

US-09-528-682-1 x BG859250 ..

Align seg 1/1 to: BG859250 from: 1 to: 686

```

84 TyrIleTyrValIleAlaThrAlaProAsnmet..... 94
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57 CATTGTGGGTTTCACAGCCCGAGAGTGGCCACCTGCTGCTCT 106
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 .....PheasnValAsnAspVal...LeuglyValTyrSerProH 107
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 CAACCTGTGGCCATCACATTCCTCAAGTGGGCTAGGCGTGTACAACTCG 156
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 lAsProTyr.....GluIngluValSerAlaLeuglygly 118
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
157 ACCCGTGGAGAGAACGTACGCGCCGCTAGCTGAGCGCTGGGCTGG 206
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
119 lLePro.....Tyr.SerGlnIleTyrGly...TyrTyrArgVala 131
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 ATACCCGCCAGCTTGATACGACAGGTGTACGCAATGAGGCCGATGTGG 256
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 snPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyr..... 145
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
257 GTGGGGCGGTGGCGGACAGGCGCTGCCGGGACCAAGGTATACGAGAC 306
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 .....ArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAl 158
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 AGCAACCTGGCGGACAGCGGTATGACAGGCGCTGGGCGCTGT 356
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 actlAspGlyTyrArgLeuAlaGlyPhePro..... 168
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
357 ACAGGACAGTATACGCGGCGCTGCCGTAACCTGATCTGCTACTGC 406
   |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

307 CGGAGACTCTCTGGCGGTCTTCCCAAGGCCCCCTGCGGGGGCTCCCTCAT 356
62 rSerLeu.....SerLeuArgS 68
   ::|||
   :|||:
357 GCGCGTGGAGCAGACGACCGTCCGCCGCGCTTGCAGCACTCCAGTCCG 406
68 eTrAlaHisLeuAlaGlyInSerTLeuSerGlyTyrSerThrTyrTyr 84
   :::::
   :|||:
407 CGGTGGGCTACAGCGCGGCTCGCTCCAGTCGGGCGGACAGCTAC... 453
85 lLeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 101
454 .....ACCACTGACCCCGCTGCGAGCCCTTGACAGCAAGTT 491
   ::|||
   :|||:
101 uGlyValTyrSerProHisProTyrGluGlnValSerAlaLeuGly 118
   :
   :|||:
492 CCACACACACACACCCCGCCACACACACACACCCCTGC..... 528
118 lYlleProTyrSerGlnIleTyrGlyTrrTyrAlaValAsnPheGlyVal 134
   ::|||
   :|||:
529 ..CTGCCCTCAGCAACGTGATCGGACGCTTCAACCTG.....ATG 567
135 lLeAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 151
   ::|||
   :|||:
568 CGCGAGAGACGGGCGGACCTGGGCGACCAACTTCTACACCCCTACGGGA 617
151 gAsnLeu.....AsnIleAlaProAlaGluAspGlyTyrArgL 164
   ::|||
   :|||:
618 GGCAGCTGGCCATGTCACGAGCCTGTCGCCCGCTCACGGGGCGGGCC 667
164 eAlaAlaGlyPheProProAspHis.GlnAlaTrrArgGluGluProTrr 180
   ::|||:
   :|||:
668 TGGCGTCCCTCGCTGACGGCTACGGCAGCCTGGGCGACACCCCAACGCG 717
180 eHisHisAlaProGlnGlyCysGlyAsnSerArgTrrIleThrGly 196
718 AACGCGACGC.....AGATGCTCACGGCGGCTACGACGCTCACGGG 760

seq_name: gb_est2:B1546037

seq_documentation_block:
LOCUS B1546037 931 bp mRNA linear EST_05-SEP-2001
DEFINITION 60318817BP1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5259786 5',
RNA sequence.
ACCESSION B1546037
VERSION B1546037.1 GI:15433349
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 931)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shireki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LHAM1654 row: k column: 19
High quality sequence stop: 778.
Location/Qualifiers
1..931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5259786"
/clone_id="NIH_MGC_95"
FEATURES
source

```


/tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; site_2: SalI-XhoI (gtagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.5 kb and
 normalized to ROP 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIMH/NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."
 BASE COUNT 181 a 250 c 289 g 211 t
 ORIGIN

alignment_scores:
 Quality: 85.00 Length: 199
 Ratio: 0.817 Gaps: 11
 Percent Similarity: 52.261 Percent Identity: 23.618

alignment_block:
 US-09-528-682-1 x B1546037 ..

Align seg 1/1 to: B1546037 from: 1 to: 931

45 AAlaArgGlyThrGlnThrGlyPheValArgTyrAspAspGlyTyrValSe 61
 127 GCACGAGAGCTGAGATGAGACACCTGGGGCAGCCGATACCCGGGGCAG 176
 61 rThrSerLeuSerLeuArgSerAlaHisLeu.....AlaGlyG 74
 177 GGCAGCCCTGGAGCAACCCCTACCCACCGAGGGGTGGTGGTGGG 226
 74 InSer.....IleLeuSerGlyTyrSerThrTyrIleTyrValIle 88
 227 GCACCTCATGTGTGGTGGGGGCGAG.....CTCAGTGTGTG 267
 89 AAlaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSe 105
 268 GCTGGGGGGAGCCCATGGTGGTGGGGGAGCAGCCCTCATGGTGGTGGC 317
 105 rProHisProTyrGlnGlnGluValSerAlaLeuGlyGlyIleProTyrS 122
 318 TGGG.....GACAGTGTGGGGGGCCCT..... 339
 122 eGlnIleTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArg 138
 340GGCGCTACGCTGTGGAGAGTGCATGACGAGCCCATC 378
 139 LeuHisArgAsnArgGluTyrTyrArgAspArgTyrTyrArg...AsnLeuAs 154
 379 ATACATTTCGCGAGTGCATGAGAGCCCTTACTATGTAACAAACATGCA 428
 154 nIleAlaProAlaGlnAspGlyTyrArgLeuAlaGlyPheProProAspH 171
 429 CCCTTACCCCAACCAAGTGTACTACAG.....CCCATGGATG 466
 171 IsGlnAlaTyrPargGlnGluProTyrIleHisHisAlaProGlnGlyCys 187
 467 AGTACAGC...AACCAAGAACACTTGTGCAC.....GACTGC 501
 188 GlyAsn.....SerSerArgThrIleThrGlyAspThrCysAsnG1 201
 502 GTCAATATCATCAATCAAGCAGCAGCGTGTACACAAACCAAGGGGGA 551
 201 uGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerIysV 218
 552 G.....AACTTACCGAGACCGCCTTAAGATGATGATGAGCGGTGG 592
 218 AluYsArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 233
 593 TTGAGCAGATGTATACACCCAGTACAGAGGAATCTCAGCGCTAT 639

seq_name: gb_gss:CNS06VRH

seq_documentation_block:

LOCUS CNS06VRH 954 bp DNA linear GSS 06-JUL-2001
 DEFINITION T7 end of clone AX0AA024E08 of library AX0AA from strain CBS 7064
 of *Pichia farinosa*, genomic survey sequence.

ACCESSION AL417507
 VERSION AL417507.1 GI:12199280

KEYWORDS GSS.
 SOURCE *Pichia farinosa*.

ORGANISM *Pichia farinosa*.

REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; *Pichia*.

AUTHORS 1 (bases 1 to 954)
 de Montigny,J., Spohner,C., Souciet,J., Tekala,F., Dujon,B.,

TITLE Wincker,P., Artiguenave,F. and Potier,S.
 Genomic exploration of the hemiascomycetous yeasts: 15. *Pichia*

JOURNAL sorbitophila
 FEBS Lett. 487 (1), 87-90 (2000)

REFERENCE 20584725
 2 (bases 1 to 954)

AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
 Botochin,Fukuhara,M., Bon,E., Brothier,P., Casaregola,S.,

de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Lorente,B.,
 Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,

Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
 Wincker,P. and Weissenbach,J.

TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

REFERENCE 20584711
 3 (bases 1 to 954)

AUTHORS Genoscope.

TITLE Direct Submission

COMMENT Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

This GSS is part of a random genomic sequencing program of thirteen
 yeast species: *Saccharomyces bayanus* var. *uvarum*, *Saccharomyces*

exiguus, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,
Saccharomyces kluyveri, *Kluyveromyces marxianus* var. *marxianus*, *Kluyveromyces*

lactis var. *lactis*, *Kluyveromyces fragilis* var. *fragilis*, *Pichia sorbitophila*,
angusta, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,
Candida tropicalis and *Yarrowia lipolytica*. Genomic inserts of 3 to

5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of

the other extremity of this insert.

FEATURES
 Location/Qualifiers

1..954
 /organism="Pichia farinosa"

/strain="CBS 7064"

/db_xref="taxon:4920"

/clone="AX0AA024E08"

/clone_11b="AX0AA"

/note="end : 77"

BASE COUNT 200 a 209 c 216 g 315 t 14 others

alignment_scores:
 Quality: 85.00 Length: 138
 Ratio: 1.149 Gaps: 7
 Percent Similarity: 53.623 Percent Identity: 26.087

alignment_block:
 US-09-528-682-1 x CNS06VRH ..

Align seg 1/1 to: CNS06VRH from: 1 to: 954

50 ThGlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerIle 66
 477 ACGGGTTCGTC...TACACTGAATGTATGTGATGTGAGCCCATCTT 523

[illegible]

```

seq_name: gb_est2:BE972508
seq_documentation_block:
LOCUS      BE972508                1040 bp      mRNA
DEFINITION 601651870F01 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:3935368 5',
            mRNA sequence.
ACCESSION  BE972508
VERSION    BE972508.1 GI:10585844
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1040)
            NIH-MGC http://mgc.ncl.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS    Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-remail.nih.gov
COMMENT     Tissue Procurement: CLONETECH Laboratories, Inc.
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LMNL at:
            http://image.lnl.gov
            Plate: LUCM777 row: k column: 17
            High quality sequence stop: 430.
FEATURES
     source
         1..1040
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="IMAGE:3935368"
             /clone_1lb="NIH_MGC_82"
             /lab_host="DH10B (TI phage-resistant)"
             /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
             SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattatggcc); 5' adaptor
             3' adaptors were used in cloning as follows: 5' adaptor
             sequence: 5'-CACGCGCATATGCGC-3' and 3' adaptor sequence:
             5'-ATTCTAGAGCGCCGAGCGCGCGCATG-dT(30)BN-3' (where B = A,
             C, or G and N = A, C, G, or T). Average insert size
             1.35 kb (range 0.9-4.0 kb). 14/15 clones contained
             inserts by PCR. This library was enriched for full-length
             clones and was constructed by Clontech Laboratories (Palo
             Alto, CA)."
BASE COUNT      353 a      329 c      217 g      141 t
ORIGIN
alignment_scores:
    Quality:      83.00      Length:      183
    Ratio:        0.976      Gaps:      9
    Percent Similarity: 46.448      Percent Identity: 24.590
alignment_block:
US-09-528-682-1 x BE972508
Align seg 1/1 to: BE972508 from: 1 to: 1040
36 GlnMetAsnIleAsnLeuTyraSphIs..AlaArgIlyThrIn..... 49
      :::::  ::  |||  |||  |||||  |||||
466 GACCTGACGTGCTCATCGTACCGGACATCGACGACGACAGAGAC 515
50 .ThrglyPheValArg.....TyrAspA 57
      |||||  |||  |||
516 GACAGGGGACCCACCGCGGAGGCTAACATAGAACAAACCAACCTG 565
57 spcIlyValSerThnSerIeuSerIeuArgSerAlaHisIeuAlaGly 73
      ||::|  |||||  ::  ||:::
566 AGCCCTACGCGCAACATATCAT..... 568
74 GlnSerIleIeuSerGlyTyrSerThrTyTyr.IleTyValIleAlaT 90

```


Tel: 5105595773
Fax: 5105595818

Email: oanderson@w.usda.gov

Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

Seq primer: Stratiagene SK primer.

Location/Qualifiers

SOURCE

1. 370
/organism="Tridicum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE1709.E10.119"
/clone_lib="Wheat heat stressed spike cDNA library"
/tissue_type="whole spike"
/dev_stage="Spike at 5, 10, 15 and 20 days after anthesis"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid: Site_1: EcoRI; Site_2: XhoI; Spikes at 5, 10, 15 and 20 days after anthesis were heat stressed under two conditions at Texas Tech University (D. Zhang in HT Nguyen lab): (1) at 38 C for 4 hours and (2) 5 days of cyclic treatment of 38 C for 4 hours. Total RNA and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the T7 Close lab (Choi, Close, Renton, Maltrasi) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 80 a 114 c 104 g 71 t 1 others
ORIGIN

alignment_scores:
Quality: 82.50 Length: 89
Ratio: 1.528 Gaps: 3
Percent Similarity: 60.674 Percent Identity: 28.090

alignment_block:
US-09-528-682-1 x BF429191 ..

Align seg 1/1 to: BF429191 from: 1 to: 370

151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
||||| : : : : : ||| : : : |||
90 AGAAGCTCCAGTCAATGCCAGAGAGCAGCAGAGGTGGCGCTTT 139
167 eProProAspHisGlnAlaTrpArgGluGluProTrpIleHisAlaP 184
||||| : : : : : ||| : : : |||
140 CCCAAGTAGAGAGAAAGCTTGGGGCCGCTCCCAATGGCTTGTCCACCGC 189
184 roGlnGlyCys.GlyAsnSerSerArgThrIleThrGlyAspThrCysAs 200
: : : : : ||| : : : : : ||| : : : |||
190 TCCGAGGTGTCAGAGGAGTCCGAGTCGATCTGAGAGAACCCCTC 239
200 nGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerL 217
: : : : : ||| : : : : : ||| : : : |||
240 GAACTCGGCCGAGAGATT.....GAGTTCAGGGCCG 271
217 yValIysArgGln.....IlePheSerAspTyrGlnSerGluVal 230
: : : : : ||| : : : : : ||| : : : |||
272 AGGTCTCCAGGCTCATGACATCATCATCACTCTACAGACAGCAG 321
231 AspIleTyrAsnArg 235
||||| : : : : : |||
322 GACATCTTCTCTGAG 336

seq_name: gb_sast2:BF422183

seq_documentation_block:

LOCUS BF422183 614 bp mRNA linear EST 28-NOV-2000
DEFINITION FM112.G12.D1.A003 Floral-Induced Meristem 1 (FM1) Sorghum
prolinquum cDNA, mRNA sequence.

ACCESSION

BF422183

VERSION

BF422183.1 GI:11410172

KEYWORDS

EST.

SOURCE

Sorghum prolinquum.

ORGANISM

Sorghum prolinquum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 614)
Cordonier-Pratt,M.-M., Gingle,A., Sudhan,M., Marsala,C. and Pratt
L.H.

REFERENCE

An EST database from Sorghum: floral-induced meristems

AUTHORS

Unpublished (2000)
Contact: Cordonier-Pratt MM

JOURNAL

Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

COMMENT

Tel: 706 542 1860
Fax: 706 542 1805
Email: empratt@uga.edu

SEQUENCE

Sequences have been trimmed to exclude PolyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV

High quality sequence stop: 583

POLYA-No.

Location/Qualifiers

FEATURES

1. 614
/organism="Sorghum prolinquum"
/db_xref="taxon:132711"
/clone_lib="Floral-Induced Meristem 1 (FM1)"
/note="Organ: Floral-Induced Meristems; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; mature plants were placed in a growth chamber for 15 days with 16 hr darkness and 8 hr light (flowering is induced by short-day conditions); 16 days after being returned to the greenhouse under natural long days during late April/early May, meristems were harvested The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

BASE COUNT

195 a 148 c 177 g 94 t

ORIGIN

alignment_scores:
Quality: 82.50 Length: 111
Ratio: 1.352 Gaps: 8
Percent Similarity: 54.955 Percent Identity: 32.432

alignment_block:
US-09-528-682-1 x BF422183 ..

Align seg 1/1 to: BF422183 from: 1 to: 614

127 TrpTyr.....ArgValAsnPhgGlyValIleAspGluArgLe 139
||||| : : : : : ||| : : : |||
178 TGGTACCGACAGCAGTACGAGTACCTCCGCACTTGGACAAAGA 227
139 uHisArg.....AsnArgGluTyrArgAspArgTyrT 150
: : : : : ||| : : : : : ||| : : : |||
228 CCATCGACAGAGTATGTCGGAGGGTCCGAGCAGCAGCGGGCGAGTC 277
150 yArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaG 166
||||| : : : : : ||| : : : |||
278 GAGCGGCTCTACGAGAGCGCCGAGAGAGAGCAGCAGAGCTCGGG 327
166 yPhePro...ProAspHisGlnAlaTrpArgGluGluProTrpIleHis 181
||||| : : : : : ||| : : : |||
328 GAGACCCCGCAGAGATGATGACGAGTGGAGGTGACCCGTCGAGCGC 377
182 HisAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAsp 198
||||| : : : : : ||| : : : |||


```

131 snpheiylalileasprglutyleuhsiaasgnalvrgllytvarasp 147
132 ||:|||||:||||:||||:||||:||||:||||:||||:||||: 148
206 acTGGAAAT....TCAGACACAGACGAAGACTGTGAGCTACAAAGAT 249
148 ArgTGTATArgAsnleuAsnIlealProalagIuAspIyTArgIe 164
250 .....GAGAAGCTGTCTTCTCTGTGCTGTGT...TTCACTGT 284
164 uAlagIyPheProProasp.....HisGlalalTPrpArgIug 177
285 GTTGGGCTGACACACACCTTTTCAGACACTACACTCTCCAGATCAAGAC 334
177 IuPrPrIleHsHisIala.....ProGInglYcysGlyAsn 189
335 AGAGTATTACATCCACGACGACAAATAATGCCCCCGTACTTGTGGA... 381
190 SerSerArgThrIleThrgly.....AspTh 198
382 .....AGCAGAACAGACCGAGGTCTTCTTCACCGAGGTGGCAACGTGACAC 428
198 rCySaSngIuGlUthGInAsnleuSerThrIleIeyrleuArgIuYrg 215
429 TGCATGCAAAATTTTATGCGACACCTACAGCATTTGGCTTAGGAATCCAC 478

```

Location/Qualifiers
1. .596

```

479  AAAATCCGGATCAAGGACACCAAGTACTTACATTCAGTTACTCAGGAAGT 528
230  IAspIleTyr 233
      |||||.....
529  AGATGCTTTC 538
name: gb-sss:AZ247543

```

seq. documentation block:	613 bp	DNA	linear	GSS 15-JUN-2000
LOCUS	A2247543			
DEFINITION	RPci-23-92p2.TV	RPci-23	Mus musculus genomic clone	RPci-23-92p2,
ACCESSION		DNA sequence.		
VERSION	A2247543			
KEYWORDS	A2247543.1	GI:8560746		
SOURCE	GSS.			
ORGANISM	house mouse.			
	Mus musculus			
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eulhelia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1 (bases 1 to 613)			
	Zhao, S., Nieman, W., Feldblum, T., Malek, J., Shatsman, S., Akinret			
	, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Kroll, M., de Jong, P.			
	and Fraser, C. M.			
TITLE	Mouse BAC End Sequences from Library RPci-23			
JOURNAL	Unpublished (1999)			
COMMENT	Other_GSSs: RPci-23-92p2.TV			
	Contact: Shaying Zhao			

cores:

Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Lock:

(pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page:

Seq primer: SP6
Class: BAC ends.

Source

```
/organism="Mus musculus"  
/strain="C57BL/6J"
```


[illegible]

DEFINITION	B619064 RIKEN full-length enriched, 8 days embryo Mus musculus cDNA clone 5730471122 5', mRNA sequence.
ACCESSION	B619064
VERSION	B619064.1
KEYWORDS	GI:15396735
SOURCE	EST.
ORGANISM	house mouse. Mus musculus.
REFERENCE	Mammalia: Eutheria: Rodentia: Sciurognathia: Muridae: Murinae: Mus. Embryos: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
AUTHORS	1 (bases 1 to 634) Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matshiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno,H., Fukuishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shingawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y. Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Funct. Genomics 2 pre, L72-L86 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.
FEATURES	Location/Organism
SOURCE	1..634 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="5730471122" /clone_id="RIKEN full-length enriched, 8 days embryo" /sex="mixed" /dev_stage="8 days embryo" /lab_host="DH10B"
	/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGGAGAAGAGATCCACAGACCTTTTTTTTTTTTTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by

cap-trapper. cDNA went through one round of subtraction to
 Rot = 100.0 second strand cDNA was prepared with the
 primer adapter of sequence 15'

GAGAGAGATTCTCGAGTATTAATTAATATCCCCCCCCC 3'. cDNA
 was cloned into the XhoI and BamHI sites. Vector: a
 modified plasmid pUC19 (+) after bulk excision from Lambda
 FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

BASE COUNT 182 a 170 c 147 g 135 t
 ORIGIN

alignment_scores:
 Quality: 82.00 Length: 154
 Ratio: 1.000 Gaps: 11
 Percent Similarity: 53.247 Percent Identity: 25.974

alignment_block:
 US-09-528-682-1 x BB619064 ..

Align seg 1/1 to: BB619064 from: 1 to: 634

```

106 ProHisProTyr.....GIUGLGLVAlSeraLaLeuGlyGly..... 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 CCACACCCACGACCGCAGCTCCGAGCAACAATTAGCTCAGCAGCAGC 102
119 .....IleProTyrSerGlnIleTyrGlyTyrTyrGlyAla 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 ACTAGCTCACTTCATCTCCAGCTCCGAGCTAAGTGAGCTCAGCTTGA 152
131 snPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArgAsp 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 ACTGGAT.....TCAGGACACGACGAGAGAGAGTCTGAGCTACAGAGAT 196
148 ArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLe 164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
197 .....GAGAGCTCTGCTTCTCTGCTGCTGAT...TTCAGTCT 231
164 uAlaGlyPheProProAsp.....HisGlnIleTyrArgGlu 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 GTTGGCTGACACCCACCTTTCAGACAGCTACCTCCAGATCAAGCA 281
177 LuProTyrIleHisAla.....ProGlnGlyCysGlyAsn 189
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 AGAGTTTTCATCTCCAGCAAGCAAGAAATGCCGCCGCTACTGTGGA... 328
190 SerSerArgThrIleThrGly.....AspThr 198
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
329 ..AGCAGACAAAGCCAAAGCTCTCTCTCACCAGAGTGGAAGCTGACAC 375
198 rCysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrG 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
376 TGGCATGCAATTTTATCGAGACCTTACCATTTGGCTCAGGAAATCCAC 425
215 InSerLysValLys...ArgGlnIlePheSerAspTyrGlnSerGluVa 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
426 AATATCGCATCAAGTGAGCAAGCACTTCACTTCACTTCACTGAGGAGT 475
230 LaSPIleTyr 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
476 AGATGCTCTTC 485
```

seq_name: gb_gss:A2571452

seq_documentation_block:
 LOCUS A2571452 687 bp DNA linear GSS 15-MAY-2001
 DEFINITION 287PVH02 PV MBN #30 Plasmodium vivax genomic 3', DNA sequence.
 ACCESSION A2571452
 VERSION A2571452.1 GI:13983551
 KEYWORDS GSS.
 SOURCE malaria parasite P. vivax.
 ORGANISM Plasmodium vivax
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1 (bases 1 to 687)

AUTHORS Carlton, J.M.-R. and Dame, J.B.
 TITLE The Plasmodium vivax and P. berghei gene sequence tag projects
 JOURNAL Parasitol. Today 16 (10), 409 (2000)

COMMENT
 CONTACT: Dame JB
 Dept. of Pathobiology, College of Veterinary Medicine
 University of Florida
 2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
 Tel: 352 392 4700
 Fax: 352 392 9704
 Email: damej@mail.ufl.edu
 Seq primer: M13(-20) forward
 Class: Shotgun.

FEATURES
 source Location/Qualifiers

```

1..687
/organism="Plasmodium vivax"
/strain="Salvador I (Collins, W. 1972. J. Parasitol. 69,
497-598)"
/db_xref="taxon:5855"
/clone_lib="PV MBN #30"
/dev_stage="asexual blood forms"
/lab_host="Salimuri boliviensis"
/note="Vector: pBluescript SK(+) vector DNA, phagemid
excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V;
Host leukocytes were extracted from P. vivax infected
blood using the following methods: first, infected blood
was activated by the addition of 0.5 ml of ADP (40mg/ml)
per 10 ml blood. Then blood was passed over a column of
acid washed 0.1 mm glass beads, then through a Plasmidpur
filter, followed by passage through a column of pre-wet
Whitman CFI powder (1:2 ratio volume of blood to CFI),
and finally centrifuged through a 50% Percoll density
cushion. Purified DNA was digested with mung bean nuclease
in the presence of 44% formamide at 50°C, as described
(Vernick, K.D., Imbericki, R.B., and McCutchan, T.F. 1988.
Nucleic Acids Research 16:6883-6896). Digested DNA was
blunt-ended using T4 DNA Polymerase and size fractionated
over a Sepharose CL-2B column. Fractions in the size range
500bp-4kb were ligated into the Eco RV site of pBluescript
SK(+), and E. coli XL-10 Gold transformed with the
ligation mixture."
```

BASE COUNT 178 a 169 c 168 g 170 t 2 others
 ORIGIN

alignment_scores:
 Quality: 82.00 Length: 206
 Ratio: 0.837 Gaps: 13
 Percent Similarity: 47.573 Percent Identity: 24.757

alignment_block:
 US-09-528-682-1 x A2571452 ..

Align seg 1/1 to: A2571452 from: 1 to: 687

```

16 IleLysArgSerGlyGlyLeuMetProArgGly.....HisAsnG1 29
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
16 TTTAAACGAGAGGGGGGGTACCCCGAGAGATTATCACAAAAAAA 65
29 uTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAla 46
: ||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 AACATTCAACAGGGGGGCC.....ATTCAACAGAGGGGCCATTCA 106
46 rGlyThr.....GlnThrGlyPheValArgTyr 55
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
107 ACGGGCGGTATTCATTAATGGGGCCCCCCCCCAAGCAAC.....CGGTAT 150
56 AspAsp.....GlyTyrValSerThrSerLeuSerLeu... 66
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 CCAACGACACCGCATCAACAGCAACCGCGCTCAAGCGCGCTGTGGG 200
67 .....ArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyT 80
||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 GAGCAACATCGTTTGGCGGCGCTGAGCGCAGAGAGCTCTGCTGTGGG 250
```


High quality sequence stop: 19.
Location/Qualifiers

FEATURES
source
1. 1935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4121047"
/clone_lib="NIH.MGC-17"
/tissue_type="rhodomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using Zap-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT
526 a 573 c 471 g 360 t 5 others

ORIGIN

alignment_scores:
Quality: 82.00 Length: 92
Ratio: 1.577 Gaps: 5
Percent Similarity: 56.522 Percent Identity: 29.348

alignment_block:
US-09-528-682-1 x BF304330 ..

Align seg 1/1 to: BF304330 from: 1 to: 1935

```

6 TTTATG.....AlaSPSeIARpProPROMAspGluIleLysArgSerGI 20
|||||.....|
1286 TACCGCATACCCGCTGACGCCACAGTTCGCCCTACTGTCAGCGCTTATGG 1335
20 yGlyLeu.....MetProArg 26
|||||.....|
1336 TGGCCCTACTTTGCGCTACCTCATATAGTTAGTACCCACACTGCTGCC. 1384
26 LYLHLSAnsluTYrPheAspArgGlyThrGlnMetAsnIleAsnLeuTYr 42
|||||.....|
1385 ..CATGCGAATTGGATGACAAAGAACCCCTTGTGATA.....CTC 1426
43 AsPnLSAlArGlyThrGlnThrGlyPheValArgTYrAspAspGlyTYr 59
|||||.....|
1427 GACGACAAACGCTAGACACACAGGAGGTTCCACAAACCCCGATGGAAA 1476
59 rVAlSerThrSerLeuSer.....LeuArgSerAlaHisLeuAlaGly 74
|||||.....|
1477 CATGCCGACACAGGAGTGCATATTAACAAGTCACACACACACAGCAGCA 1526
74 InSerIleLeuSerGlyTYrSerThr 82
|||||.....|
1527 CCGCATGTGCGACGACACAGAAC 1552

```

seq_name: gb_est2:BE864213

seq_documentation_block:
LOCUS BE864213 368 bp mRNA linear EST 29-SEP-2000
DEFINITION UI-M-BH1-ant-c-02-0-UI.r1 NIH.BMAP.M.S2 Mus musculus cDNA clone
UI-M-BH1-ant-c-02-0-UI 5', mRNA sequence.
ACCESSION BE864213
VERSION BE864213 GI:10385033
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 368)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE
COMMENT
9704447
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@mail.nih.gov
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements
Seq primer: M13 Reverse.

FEATURES
source
1. 368
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH1-ant-c-02-0-UI"
/clone_lib="NIH.BMAP.M.S2"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The NIH.BMAP.M.S2 library is a subtracted library derived from NIH.BMAP.M.S1, which in turn is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 5,000 clones from the NIH.BMAP.M.S1 library and a pool of 2,000 clones obtained from non-normalized and normalized mouse brain spinal cord libraries."

BASE COUNT
83 a 98 c 94 g 93 t

ORIGIN

alignment_scores:
Quality: 81.50 Length: 132
Ratio: 1.273 Gaps: 8
Percent Similarity: 48.485 Percent Identity: 25.758

alignment_block:
US-09-528-682-1 x BE864213 ..

Align seg 1/1 to: BE864213 from: 1 to: 368

```

101 LeuGlyValTYrSerProHisProTYrGlu.....GlnGluValSe 114
|||||.....|
10 CTTCTGTGTCTTCACAAACCCCAATCCAGGAATCTTCAGCATGCTCA 59
114 rAlaLeuGlyGlyIleProTYrSerGlnIleTYrGly...TPTTYr...A 129
|||||.....|
60 CACTGTGAATGGATTACCTACATGCTGACAGCGACAGAGACATCTTA 209
129 rGValAsnPhGlyValIleAspGluArgLeu..... 139
|||||.....|
110 GGAAGCTGTGGGGCGTGGCTTCCTCCGCTAAGTGCAGCTGTCACAT 159
140 .....HisArgAsnArgGlyTYr.....ArgAspArgTYr 149
|||||.....|
160 GCCTACGTCATCTACATGCTGACAGCGACAGAGACATCTTA 209
149 rTYrArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaG 166
|||||.....|
210 CTTT.....GCACACCGTGTGCC. 228
166 LYPheProPROMAspHisGlnAlaTPArgGluInuProTYrIleHis 182
|||||.....|
229 .....CAAGCTTGGGAACGGCCGACGTTCAATCCACAC 261

```


[illegible]

296	ATCTTCACGAGAAACCACCTCGCCGAACCTCATTGAAAGAATCTTTTAACGT	247
97	IAsnaSpvAlIeuglyValTyrSerProHisProTyrgluIngluValS	114
246	CCTCTCGGGCTTCCATATGACAGCCACCACTGACTGAAATGGTGCT	197
114	eValaleu.....glygylleprotyrserglinlelyrgltyrpyr	128
196	CGCTCACATTCAGTGGGAATVGAACCTCTTCCGAG.....TTTAT	156
129	ArgeValasPheglYval.IleaspLuarlgreuhHsarGasnArgglut	145
155	CGTCTCATTTGCCGATAGCTGAGACAATCACTTCTATCAAGAGAAATCTC	106
145	yfArgasp	147
105	TGAGGCAT	98
seq_name:	gb_est2:Bj187033	
seq_documentation_block:		
LOCUS	Bj187033	608 bp mRNA linear EST 24-JAN-2002
DEFINITION	Bj187033 normalized full length cDNA library, chloronemata, caulonemata and malformed buds Physcomitrella patens subsp. patens cDNA clone pphb39116 5' , mRNA sequence.	
ACCESSION	Bj187033	
VERSION	Bj187033.1	GI:18354974
KEYWORDS	EST.	
SOURCE	Physcomitrella patens subsp. patens.	
ORGANISM	Physcomitrella patens subsp. patens. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella. 1 (bases 1 to 608) Fujita,T., Shln-i,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T, , Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe ,M.	
REFERENCE		
AUTHORS		
JOURNAL	Comparison of the moss Physcomitrella patens genome with flowering plants genome unpublished (2002)	
COMMENT	Contact: Tadasu Shln-1 Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshln@genes.nig.ac.jp A backbone of the vector is pluscript II, that was in vivo excised from a modified IPS phage vector (Mo bi Tec, Germany). XhoI digested-5' end of cDNA is ligated to SalI site of the vector, and the BamHI digested-3' end including poly-A tail is ligated to BamHI site of the vector. cDNA insert could be amplified with conventional T7 and T3 primers. This normalized full-length cDNA library was generated basically according to the method described in Genome Research 10, 1617-1630 (2000). Carninci, P. et al. Protonemata were blended by the POLYTIRON, and then cultivated on the BCD medium containing 0.5m BA (benzylaminopurine) for 8 to 13 days under the continuous light.	
FEATURES		
source	location/Qualifiers	
	1..608	
	/organism="Physcomitrella patens subsp. patens"	
	/db_xref="taxon:145481"	
	/clone_pphb39116"	
	/clone_lib="normalized full length cDNA library, chloronemata, caulonemata and malformed buds"	
	/tissue_type="mixture of chloronemata, caulonemata and malformed buds"	
BASE COUNT	156 a 158 c 169 g 125 t	
ORIGIN		
alignment_scores:	Quality: 81.50 Length: 70	
	Ratio: 1.772 Gaps: 4	

DEFINITION 602759363F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4895030 5',
 mRNA sequence.
 ACCESSION B1199325
 VERSION B1199325.1 GI:14654346
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 807)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@rs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
 Plate: L10M1779 row: m column: 15
 High quality sequence stop: 798.
 Location/Qualifiers
 1..807
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4895030"
 /clone_lib="NIH_MGC_19"
 /tissue_type="neuroblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 182 a 232 c 213 g 180 t
 ORIGIN
 alignment_scores:
 Quality: 81.00 Length: 154
 Ratio: 1.095 Gaps: 8
 Percent Similarity: 48.052 Percent Identity: 25.974
 alignment_block:
 US-09-528-682-1 x B1199325/rev ..
 Align seg 1/1 to reverse of: B1199325 from: 1 to: 807
 61 SerThSerLeuSerLeuArgSerAlaHisLeuAlaGlyInSerIleLe 77
 :::::::::::::::::::: ::::::::::::::::::::
 487 GCACATGCGTATGTCAGTGCATCCATCTTCGTCACGAAT 443
 :::::::::::::::::::: ::::::::::::::::::::
 77 userGlyTyrSer.....ThrTyrTyrI 85
 :::::::::::::::::::: ::::::::::::::::::::
 442 .TCTGGGTACGGACACGAGGTCAAATCTGTGGGATTTCTCATTTGCC 394
 :::::::::::::::::::: ::::::::::::::::::::
 85 leTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspValLeu 101
 :::::::::::::::::::: ::::::::::::::::::::
 393 CTATATCTCCGTCGTCACCATCT..TATAGTGTCCACAGATGTGG 347
 :::::::::::::::::::: ::::::::::::::::::::
 102 GlyValTyrSerProHisProTyrGluGlnGlnValSerAlaLeuGly.. 117
 :::::::::::::::::::: ::::::::::::::::::::
 346 GGCCCGACAGATCTTCATCCACGCGCACCTGCTAGCATGTCGTC 297
 :::::::::::::::::::: ::::::::::::::::::::
 118GlyIleProTyrSerGlnI 124
 :::::::::::::::::::: ::::::::::::::::::::
 296 ATCTCATCCACCCACCCCATTTATTCATTGGGGAAGCCATTGATC.... 251

124 leTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArgLeuHis 140
 :::::::::::::::::::: ::::::::::::::::::::
 250TTCAGTACTGGTCTCAGAGTAAAGTCTGAGACTCTCCG 212
 :::::::::::::::::::: ::::::::::::::::::::
 141 ArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaPr 157
 :::::::::::::::::::: ::::::::::::::::::::
 211 AAGTACTGGGGGTAC...GGAGGCTGTATCCAAACTGTTCATATCCAAC 165
 :::::::::::::::::::: ::::::::::::::::::::
 157 oHa.....GluAspGlyTyrA 163
 :::::::::::::::::::: ::::::::::::::::::::
 164 GGCACATCGGGGGTCCCGGGGGTCCACACATACAGATTTGTGTCATT 115
 :::::::::::::::::::: ::::::::::::::::::::
 163 rGleuAlaGlyPheProPheAsnHisGlnAlaThrArgGluGlnProTyr 179
 :::::::::::::::::::: ::::::::::::::::::::
 114 TTCTGGCAGAGGTCCA.....CATCGTCAGAGAACAGGCAATCCCACT 71
 :::::::::::::::::::: ::::::::::::::::::::
 180 lIeHisHisAla 183
 :::::::::::::::::::: ::::::::::::::::::::
 70 CTCATCACGCA 59
 seq_name: gb_est2:B1152296
 seq_documentation_block:
 LOCUS B1152296 1401 bp mRNA linear EST 05-JUL-2001
 DEFINITION 602917727F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5068039 5',
 mRNA sequence.
 ACCESSION B1152296
 VERSION B1152296.1 GI:14612297
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1401)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@rs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L10M1183 row: f column: 08
 High quality sequence stop: 153.
 Location/Qualifiers
 1..1401
 /organism="Mus musculus"
 /strain="C57BL/6J (f1)"
 /db_xref="taxon:10090"
 /clone="IMAGE:5068039"
 /clone_lib="NCI_CGAP_Lu29"
 /tissue_type="spontaneous tumor, metastatic to mammary.
 stem cell origin."
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pCMV-Sport6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 466 a 568 c 246 g 121 t
 ORIGIN
 alignment_scores:
 Quality: 81.00 Length: 208
 Ratio: 0.818 Gaps: 11
 Percent Similarity: 47.596 Percent Identity: 24.038
 alignment_block:
 US-09-528-682-1 x B1152296 ..

Align seg 1/1 to: BF152296 from: 1 to: 1401

```

22 LeuMePrOaRgLyHIsAsnGluTyRPhAsPaRgLyThrClnMetas 38
36 CTAACTTCAAGACCAAGAAATGGGACACAGCAAGAAAGCTCATCA 85
38 nileAsnLeuTyRAspHis.....AlaArgGlyThrGlnThrGlyP 52
86 CATCGTCGTATATCGGACACGTAGATCCGGCAAGTCACACCAACGGCC 135
52 heValArgTyRAspAspGlyTyRValSerThr.SerLeuSerLeuArgS 68
136 ACCTGATCTCAAAATGTGGCTGGACTCGACAGACGCAACCACTCGAAA 185
68 rAlaHISLeuAlaGlyGlnSerIleLeuSer.....GlyTyRSerT 82
186 CCGTTCGCAAGCAACGACCGCTGCTGAGCATGGCAAGCGGCATCTT 235
82 hTyTyTyRLeuTyRValIleAlaThrAlaProAsnMetPheAsnVala 98
236 CAAGATACGCGCTG...CGATCTTCAGACCAAACTGCAACAGCTGAGC 282
99 AspValLeuGlyValTyRSerProHISProTyRGlUGlnValSerAl 115
283 CCGCTG.....CACCGCTGCCA..... 300
115 aLeuGlyGlyIleProTyRSerGlnIleTyRGlTyR..... 128
301 .....CTGACCTACTCGACATCTTACCCCTGCGCAACCTCGC 340
129 .....ArgValAsnPhGlyValIle..... 135
341 AGCACCAGCCCAACACTACTCTACTCTGACCCACTTTCACATGCC 390
136 .....AspGluArgLeuHISArgAsnArgGluTyRArgAspAr 148
391 CCCAGCGCAACACGACGACCTTTCATCCCAACAAACATGACCAAC 440
148 gTyTyTyRArgAsnLeuAsnIleAlaProAlaGluAspGlyTyRArg 165
441 A.....GGCCACCCCTCCCAACCCGCTCGACC 469
165 laGlyPheProProAsnHISGlnAlaIleTPArgGluLupProTr...Ile 180
470 CAGGCTCCCGCCGACCTGCACTGACGACGACGACGACGACGACGAC 519
181 HISAlaIleProGlnGlyCysGlyAsnSerSerArgThrIleThrGly 197
520 CACGACGACACCTTAC.....GGCACAAGCCCAAGAGCTATCATCC 563
197 pThrCysAsnGluGluThrGln 204
564 GAACCAAAACCGAGACAAAG 585

```

seq_name: gb_est2:BF345560

seq_documentation_block:

```

LOCUS      BF345560                1576 bp      mRNA      linear      EST 22-NOV-2000
DEFINITION 602019145F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4154602
5' mRNA sequence.
ACCESSION  BF345560
VERSION     BF345560.1 GI:11293155
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 1576)
AUTHORS   NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.

```

Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9424 row: 3 column: 11
High quality sequence start: 3
High quality sequence stop: 299.
Location/Qualifiers
1. 1576
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:4154602"
/clone_id="NCI_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="TDH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP library."

BASE COUNT 592 a 472 c 375 g 137 t
ORIGIN

alignment_scores:
Quality: 81.00 Length: 150
Ratio: 1.052 Gaps: 5
Percent Similarity: 51.333 Percent Identity: 22.000

alignment_block:
US-09-528-682-1 x BF345560 ..

Align seg 1/1 to: BF345560 from: 1 to: 1576

```

79 GlyTyRSerThrTyTyTyRValIleAlaThrAlaPro..... 92
11 ..... 111
138 GGAATGTAATCGTATGCTCGATATCTGAGACGGCATCTCTTGTG 187
93 .....AsnMetPheAsnValAsnAspValI 101
188 AGACCCCATTTGCGTGATGCACTAACGGTGATGACCATTCATTCATGC 237
101 euGlyValTyRSerProHISProTyRGlUGlnValSerAlaIleuGly 117
238 TCGGACATGAGGATATATCATATAGAAATTCACAGAAAGGCGAG 287
118 GlyIleProTyRSerGlnIleTyRGlTyTyRArgValAsnPhGlyVa 134
288 GAAGACCCCTTACTACTGACACAAAGG.....ATCAGCATGAAGAT 328
134 lIleAsp.....GluArgLeuHISArgAsnArgGluTyR 146
329 TACAGATACACAGAGAAACTAGATAGAGAGCGGTACAAAAGATAGC 378
146 rGAspArgTyTyTyRArgAsnLeuAsnIleAlaProAlaGluAspGlyTy 162
379 CCGTGCATGGCCAGAAATAGAACATCCGGAACATGAGACGTACTAC 427
163 ArgLeuAlaGlyPheProProAsnHISGlnAlaIleTPArgGluLup 179
428 AGACTACAGGTGACCCACG..... 448
179 pIleHISAlaIleProGlnGlyCysGlyAsnSerSerArgThrIleThr 196
449 .....CACCGACATATATGC.....ACGCGACAGACGACCAATG 482
196 lAspPhrCysAsnGluGluThrGlnAsnLeuSerThrIleTyRLeuArg 212
483 GACCACGCGCAAGCGACACACAGAAAGACAGACTGCGTACATAAAGA 532

```



```

seq_name: gb_est2:BF520534
seq_documentation_block:
  LOCUS      BF520534          464 bp      mRNA      linear      EST 08-DEC-2000
  DEFINITION  EST458006 D5IL Medicago truncatula cDNA clone pdsil-24M2, mRNA
  sequence.
  ACCESSION  BF520534
  VERSION    BF520534.1 GI:11609217
  KEYWORDS   EST.
  SOURCE      barrel medic.
  ORGANISM   Medicago truncatula
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
              Medicago.
  REFERENCE  1 (bases 1 to 464)
              Fedorova,M., Pleeson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
              H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
              Holt,I.E. and Fraser,C.M.
              ESTs from leaves of Medicago truncatula after inoculation with
              Colletotrichum trifolii
  JOURNAL    Unpublished (2000)
  COMMENT    Contact: Deborah A. Samac
              Department of Plant Pathology
              University of Minnesota
              495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
              Tel: 612 625 1243
              Fax: 651 649 5038
              Email: debbyspuccln1.crl.umn.edu
              University of Minnesota name: M277975e
              TIGR sequence name: MTECV73TK
              More information is available at: http://chrysie.tamu.edu/medicago
              Seq primer: SKmod (CTA GAA CTA gta ggt CC).
  FEATURES
    source
      1..464
        /organism="Medicago truncatula"
        /cultivar="genotype A17"
        /db_xref="taxon:3880"
        /clone="pdsil-24M2"
        /clone_lib="D5IL"
        /tissue_type="Leaves infected with Colletotrichum
        trifolii"
        /dev_stage="cotyledons and primary leaves harvested 5 and
        8 days after inoculation with Colletotrichum trifolii"
        /lab_host="E. coli strain XLOLR"
        /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
        XhoI; cDNA was prepared from polyA+ enriched RNA from
        cotyledons and primary leaves harvested 5 and 8 days after
        inoculation with Colletotrichum trifolii. The cDNA was
        directionally ligated into the Uni-ZAP XR vector from
        Stratagene and packaged using Gigapack III gold packaging
        extracts. Plasmids containing cDNA inserts were excised
        from the recombinant lambda-ZAP phage using Ex Assist
        helper phage and propagated in XLOLR cells. Note: EST may
        be of fungal origin."
  BASE COUNT      124 a      101 c      88 g      151 t
  ORIGIN
alignment_scores:
  Quality:      80.50      Length:      72
  Ratio:        1.750      Gaps:      3
  Percent Similarity: 63.889      Percent Identity: 33.333
alignment_block:
  US-09-528-682-1 x BF520534/rev ..
  Align seg 1/1 to reverse of: BF520534 from: 1 to: 464
      2 GYAsPARGLeuTYrARGAlaAspSerArgProProAspGluLeuysAr 18
      |||:::|||||:::|||||:::|||||:::|||||:::
      412 GGACAAAGTTCTATGTAGTC.....CCATAAGACACATACGAAA 372

```

```

18 gSerGlyLeuMeProArgGlyHisAsnGluTYrPheAspArgGlyT 35
   :|||||:::|||||:::|||||:::|||||:::|||||:::
371 GCTTCAAGGTCCTTCTCCAGGAACTCTCT.....TCTCCTGCTC 331
35 hGclMetAsnIle...AsnLeuTYrAspHisAlaArgGlyThrGlnThr 50
   ::|||:::|||||:::|||||:::|||||:::|||||:::
330 GTGAATGACATCGGCTTCTTACGATGAGACGTAAACACACATCC 281
51 GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeuAr 67
   |||:::|||||:::|||||:::|||||:::|||||:::
280 CGCAGATGAGGAAACATGATGTATGATCTACTTCTGCTATCTTAG 231
67 gSerAlaHisLeuAla 72
   ||:::|||||:::|||||:::|||||:::|||||:::
230 GAAGAAACATATTGCG 215
seq_name: gb_est1:AU209914
seq_documentation_block:
  LOCUS      AU209914          514 bp      mRNA      linear      EST 17-JUL-2001
  DEFINITION  AU209914 unpublished oligo-capped cDNA library, stage I1
  Caenorhabditis elegans cDNA clone YK750d06 3', mRNA sequence.
  ACCESSION  AU209914
  VERSION    AU209914.1 GI:14845824
  KEYWORDS   EST.
  SOURCE      Caenorhabditis elegans.
  ORGANISM   Caenorhabditis elegans.
              Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdicoidea
              ; Rhabdilitidae; Pelodermidae; Caenorhabditis.
  REFERENCE  1 (bases 1 to 514)
              Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
              and Sugano,S.
              A complementary view of the C.elegans genome
  JOURNAL    Unpublished (2001)
  COMMENT    Contact: Yuji Kohara
              Genome Biology Lab.
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411, Japan
              Tel: 81-559-81-6854
              Fax: 81-559-81-6855
              Email: ykohara@lab.nig.ac.jp.
  FEATURES
    source
      1..514
        /organism="Caenorhabditis elegans"
        /strain="N2"
        /db_xref="taxon:6239"
        /clone="YK750d06"
        /clone_lib="unpublished oligo-capped cDNA library, stage
        I1"
        /sex="Hermaphrodite"
        /tissue_type="whole animal"
        /dev_stage="I1"
  BASE COUNT      97 a      118 c      154 g      145 t
  ORIGIN
alignment_scores:
  Quality:      80.50      Length:      137
  Ratio:        1.184      Gaps:      7
  Percent Similarity: 49.635      Percent Identity: 24.088
alignment_block:
  US-09-528-682-1 x AU209914/rev ..
  Align seg 1/1 to reverse of: AU209914 from: 1 to: 514
      102 GlyValTYrSerProHisProTYrGluGlnGluValSerAlaLeuGly 118
      |||:::|||||:::|||||:::|||||:::|||||:::
      390 GGGATTCATCATCGGTACCATCATCATCAACA.....GA 356
      118 yIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPheGlyValI 135
      ::|||:::|||||:::|||||:::|||||:::|||||:::

```



```

355 TCTACCGTATTGGAAAGCTGCTCT.....CACTGAGAAAGG 318
135 leaspgluargleuhlsatgasnarvglutyratrgaspargltyrtylrg 151
317 AAGAACCAAGGATCCACGAGTTCCTTACCACCAAGCATCCACCC 268
152 AsnleAsnllAlaAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheP 168
267 AATGGAGAGATTCGCCAAGATACGG.....TATGCAACACGAGACTA.. 226
168 oProAspHisGlnAlaTP.....ArgGluGluProTyrPheHisH 182
225 ...CATCATGCTCTGTCGAGCCGTTCTCGACCAAGAACGCTTTATC 180
182 lAlaAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThr 198
179 ACCCTTCGCAAGTCGCTCATCACCA..... 154
199 CysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrG 215
153 ...AACCAAGAGATCGCCCA.....CGAAGATATCA 125
215 nSerIysValIysArgGlnIlePheSerSpYrGlnSerGluValAspI 232
124 ACCCTCAAGAGATCGACACCACTTCCAAAGACCGGACACAGCTCGTCCA 75
232 leTyrAsnArg 235
74 ACCACACGCCCA 64
seq_name: gb_est1.A02222397

```

LOCUS	533 bp	mRNA	linear	EST 17-JUL-2001
DEFINITION	AU222397	unpublished oligo-capped cDNA library, stage L1		
ACCESSION	AU222397	Caenorhabditis elegans cDNA clone yk1015g07.3		mRNA sequence.
VERSION	AU222397.1			
KEYWORDS	EST.			
SOURCE	Caenorhabditis elegans.			
ORGANISM	Caenorhabditis elegans.			
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea.			
AUTHORS	Rhabdilitidae; Pelodertinae; Caenorhabditis.			
	1 (bases 1 to 533)			
	Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.			
	and Sugano, S.			
TITLE	A complementary view of the C.elegans genome			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Yuji Kohara			
	Genome Biology Lab.			
	National Institute of Genetics			
	Yata 1111, Mishima, Shizuoka 411, Japan			
	Tel: 81-559-81-6854			
	Fax: 81-559-81-6855			
	Email: ykohara@lab.nig.ac.jp.			
FEATURES	Location/Qualifiers			
source	1..533			
	/organism="Caenorhabditis elegans"			
	/strain="N2"			
	/db_xref="taxon:6239"			
	/clone="yk1015g07"			
	/clone_lib="unpublished oligo-capped cDNA library, stage			
	L1"			
	/sex="Hermaphrodite"			
	/tissue_type="whole animal"			
	/dev_stage="L1"			
BASE COUNT	101 a 123 c 156 g 153 t			
ORIGIN				
alignment_scores:	Quality: 80.50	Length: 137		
	Ratio: 1.184	Gaps: 7		

```
Percent Similarity: 49.635    Percent Identity: 24.088
alignment_block:
US-09-528-682-1 x AU222397/rev ..
Align seg 1/1 to reverse of: AU222397 from: 1 to: 533
```

```

102 gUAla1tYrSerPrOHIsrProtyrGluGlnValSerAlaIeuc1y1 118
118 yllleProtyrSerGlnIleTyrglyTPrTyraValAsnphgIyVal1 135
135 leAspGluArgleuHisrArgAsnArgGluTyraArgAspArgTyraG 151
151 AsnleAsnIleAlaProAlaGluAspGlyTyraArgleuAlaGlyPhe 168
168 oProAspHisGlnAlaTPr.....ArgGluGluPTrPrIleHisH 182
182 lSAlaPrGlnGlyIcysGlyAsnSerSerArgThrIleThrGlyAspTh 198
198 CysAsnGluGluThrGlnAsnLeuSerThrIleTyrlenuArgGluTy 215
215 nSerIyValysArgGlnIlePheSerAspTyrglnSerGluValAsp 232
232 letYrAsnArg 235
271 AATGGAGAGATGCCAAGATACG.....TATCGTCAACACGAGCTA.. 230
229 ...CATCATGCTTGTGTGAGCCGCTTCGACCAAGAACGCTTATAC 184
183 ACCCTCCGCACTCGCTCATCACCA..... 158
199 CysAsnGluGluThrGlnAsnLeuSerThrIleTyrlenuArgGluTy 215
157 ...AACCAAGAGATCGCCCA.....CGAGAAAGATCA 129
215 nSerIyValysArgGlnIlePheSerAspTyrglnSerGluValAsp 232
128 ACCTCAAGAGATCGACACCACTTCACAAACCGACGACCGCTGCTTCA 79
232 letYrAsnArg 235
78 ACCACAGCCGA 68

seq_name: gb_est1:AU209839

seq_documentation_block:
LOCUS AU209839 559 bp mRNA linear EST 17-JUL-2001
DEFINITION AU209839 unpublished oligo-capped cDNA library, stage I1
Caenorhabditis elegans cDNA clone YK749d06 3', mRNA sequence.
ACCESSION AU209839
VERSION AU209839.1 GI:14845691
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE 1. Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
Unpublished (2001)
JOURNAL Contact: Yuji Kohara
COMMENT Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 559
FEATURES
source
/oranism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"

```



```

/clone="yk749g06"
/cclone_lib="unpublished oligo-capped cDNA library, stage
L1"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
BASE COUNT      106 a      130 c      163 g      159 t      1 others
ORIGIN

alignment_scores:
    Quality:      80.50      length:      137
    Ratio:        1.184      Gaps:      7
Percent Similarity: 49.635      Percent Identity: 24.088

alignment_block:
US-09-528-682-1 x AU209839/rev ..

Align seg 1/1 to reverse of: AU209839 from: 1 to: 559

102 GlyValTyrSerProHisProTyrGluGluValSerAlaLeuGly1 118
||||| |||||: : : : : : : : : : : : : : : : : : :
394 GGATTCATCATCACCGACCATCATCAACA... ..CA 360

118 ylleProTyrSerGlnIleTyrGlyTrrPrrArgValAsnPhgIVal1 135
|||||: : : : : : : : : : : : : : : : : : :
359 TCACCGTATTGGAAAGTCTGCTCT... ..CACTGAGGAAGA 322

135 leAspIuArGLeuHISArGAsnArGluTyrArGAspArGlyTrrArg 151
: : : : : : : : : : : : : : : : : : : : : : : :
321 AGAACACACGATCCACCGAGTTCGATTCACCGAACACCATCACCC 272

152 AsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhe 168
||| : : : ||| : : : ||||| : : : |||: :
271 AATGGAGGATGCCAAGATACGG... ..TATCGTCAACACGAGCTA.. 230

168 oProAspHisGlnAlaTrp... ..ArgGluGluProTrpIleHis 182
||| ||| : : : ||| : : : ||| : : : |||
229 ...CATCATGCTCTCGTGGAGCCGTTCTGGACCAAGACGCTTATC 184

182 lAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThr 198
|| ||||| : : : : : : : : : : : : : : : : : :
183 ACCCTCCGCAAGTCGCTCATCACCCA... .. 158

199 CysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArGluTrrG1 215
|||: : : : : : : : : : : : : : : : : : ||
157 ...AACCAAGAGAGTCGCCCA... ..CGAAGATATCA 129

215 nSerIysValIysArgGlnIlePheSerAspTrrGlnSerGluValAsp1 232
| : : : ||| | : : : |||: : : ||| : : : : :
128 ACCTCAAGGATCGACACCGATTCACAGACCGACACGTCGTCCTCAAA 79

232 lEtyrAsnArg 235
: : : : : |||
78 ACCACAGCCGA 68

seq_name: gb_est1:AU215136
seq_documentation_block:
LOCUS AU215136 571 bp mRNA linear EST 17-JUL-2001
DEFINITION AU215136 unpublished oligo-capped cDNA library, stage L2
Caenorhabditis elegans cDNA clone yK821909 3', mRNA sequence.
ACCESSION AU215136
VERSION AU215136.1 GI:14853293
KEYWORDS EST.
SOURCE
ORGANISM Caenorhabditis elegans.
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabditoidea
; Rhabditidae; Pelodierinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 571)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
TITLE

```

JOURNAL
COMMENT

Unpublished (2001)
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES

SOURCE

Location/Qualifiers
. . .571
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk821g09"
/clone_1fb="unpublished oligo-capped cDNA library, stage L2"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L2"

BASE COUNT 107 a 135 c 167 g 160 t 2 others

ORIGIN

alignment_scores:

Quality: 80.50 length: 137
Ratio: 1.184 Gaps: 7
Percent Similarity: 49.635 Percent Identity: 24.088

alignment_block:
US-09-528-682-1 x AU215136/rev ..

Align seg 1/1 to reverse of: AU215136 from: 1 to: 571

102 GUYAATyTserProHisProTyrgIngluValSerAlaLeuglcl 118
||||| : ||||| :::::|||||
381 GGATTTCATCATCACCCTACCATCATCAACA.....GA 347
118 yTlePrOtyrSerGlnIleTyrgelyTpTyrrArvaLasnpheglyal 135
::::| | | ::::: ::::: |||
346 TCTACCGTATTGGAAAGTCTGCT.....CACTGAGAAGA 309
135 leaspoLuArgLeuHisArgAsnArgLUtyrARgspargrTyrrarg 151
::::::::::::||| ||| ::::: |||
308 AAGAACACAGATCCCGAGTTTGACTTACCACGAGACACATCACCCC 259
152 AsnLeuAsnIlEalaProlaglunspRgiTYrArgLeualaglyphepr 168
||| ::::| | | ::::: |||
258 ANTGGAGAGATCCCCAAGATACGG.....TATGTCAACCGAGACTA.. 217
168 oProAsphISGlnaLaTRP.....ArgglugluPRotrpIlleISH 182
||| ::::| | | ::::: |||
216 ...CATCATGTCTCGTGAGGCCGTTCTCGGACCAAGAAGCGTCTTATC 171
182 IsalaProlingLucysGLysnsrserserArgrThrIlethrGLysapTrn 198
||| ||||| ::::: |||
170 ACCCTCCGGAAGTGCCTCATCACCCA..... 145
199 CysasnGLUGlutThrglnAsnLeuserrThrlEtYrLeuarGLutyrcl 215
||| ::||| ::::: ||||| |||
144 ...AACCAAGAGAGTGCCECA.....CGAGACAGATCA 116
215 nSerlyValLySarGlnIlEpheSeraspTyrgInserrGLuaLasPI 232
| ::||| | ::||| ::::: |||
115 ACCTCAAGTGAGATCGACACACAGATTCCAAAGACCGGACAGCGTGCCTCAA 66
232 lETyrAsnarYg 235
:::| | |
65 ACCACAGCCGA 55

seq_name: qb_est1:AU211806

seq_documentation_block:
LOCUS AU211806 586 bp mRNA linear EST 17-JUL-2001


```

DEFINITION      AU211806 unpublished oligo-capped cDNA library, stage L2
ACCESSION       AU211806
VERSION         AU211806.1
KEYWORDS        GI:14849353
SOURCE          EST.
ORGANISM        Caenorhabditis elegans.
                Caenorhabditis elegans.
                Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
REFERENCE       1 (bases 1 to 586)
AUTHORS         Kohara,Y., Shii-1,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,I.Y.
                and Sugano,S.
TITLE           A complementary view of the C.elegans genome
JOURNAL         Unpublished (2001)
COMMENT         Contact: Yuji Kohara
                Genome Biology Lab.
                National Institute of Genetics
                Yata 1111, Mishima, Shizuoka 411, Japan
                Tel: 81-559-81-6854
                Fax: 81-559-81-6855
                Email: ykohara@lab.nig.ac.jp.
FEATURES
source         1..586
                /organism="Caenorhabditis elegans"
                /strain="N2"
                /db_xref="taxon:6239"
                /clone="YK776h04"
                /clone_1lb="unpublished oligo-capped cDNA library, stage
                L2"
                /sex="Hermaphrodite"
                /tissue_type="whole animal"
                /dev_stage="L2"
BASE COUNT     113 a 137 c 169 g 167 t
ORIGIN
alignment_scores:
                Quality: 80.50      length: 137
                Ratio: 1.184        Gaps: 7
                Percent Smlarity: 49.635      Percent Identity: 24.088
alignment_block:
US-09-528-682-1 x AU211806/rev ..
Align seg 1/1 to reverse of: AU211806 from: 1 to: 586
102 GYValTYrSerProHisProTyrGIuGIuValSerAlaLeuGIyGI 118
   |||::: |||||::: ::::|||:::
393 GGGATTCATCCACCGTACCTCATCAACAACA.....CA 359
118 yIleProTYrSerGlnIleTYrGIYrPYrArValAsnPhelGIyVal 135
   ::|||::: ::::: ::::: :::::
358 TCYACGGTATTGGAAAGTCTGCTCT.....CACTGAGGAAGA 321
135 leAspGIuArGIeHisArGAsnArGIuTYrArGAsparGIYrTYrArG 151
   ::::|||::: |||::: |||::: :::::
320 AAGAACACAGGGATCCACGCGATTCCTTACCAAGACACATCACCCC 271
152 AsnLeuAsnIleAlaProAlaGIuAaPglTYrArGIeuaAlaGIyHePr 168
   |||::: |||::: ::::: ||||| |||:::
270 AATGGAGAGCTTCCCAAGATACGG.....TATCGTACACCAAGACTA.. 229
168 oProAspHisGlnAlaTrP.....ArgGIuGIuProTrPleHisH 182
   |||::: |||::: |||::: |||:::
228 ....CATCATGCTTCGTGGAGCCGCTTCGGACCAAGAAGCCTTTATTC 183
182 tsAlaProGIuGIyCysGIyAsnSerSerArGIrThIleThrGIyAspThr 198
   ||| ||||| :::::
182 ACCCTCCGCAAGTCGCTCATCACCA..... 157
199 CysAsnGIuGIuTrGIuAsnLeuSerThrIleTYrLeuArGIuTYrGI 215
   |||::: |||::: |||::: |||:::
156 ...ACCCAAAGAGTGCACCA.....CGACAAAGATCA 128

```

```

215 nserlvslylsyrglnllephearspyrglnsergluvalaspi 232
|  ::::::::::| ::::| ::::| ::::| ::::| ::::|
127 ACCTCAAGTGGATCCGACACCACTTCCAGACCGGACGCGTCCCTTCCA 78

232 letyrasnary 235
:::|::|
77 ACCACAGCCGA 67

seq_name: gb_est1:AU221819

seq_documentation_block:
LOCUS AU221819 599 bp mRNA linear EST 17-JUL-2001
DEFINITION AU221819 unpublished oligo-capped cDNA library, stage I1
ACCESSION AU221819
VERSION AU221819.1 GI:14859976
SOURCE EST.
ORGANISM Caenorhabditis elegans.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae
1 (bases 1 to 599)
; Rhaditidae; Peloderinae; Caenorhabditis.
Kohara,Y., Sh1n-1,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.,
and Sugano,S.
A complementary view of the C.elegans genome
Unpublished (2001)
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1..599
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk898e07"
/cgene_libs="unpublished oligo-capped cDNA library, stage
I1"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

BASE COUNT 117 a 142 c 172 g 168 t
ORIGIN

alignment_scores:
Quality: 80.50 Length: 137
Ratio: 1.184 Gaps: 7
Percent Similarity: 49.635 Percent Identity: 24.088

alignment_block:
US-09-528-682-1 x AU221819/rev ..
Align seg 1/1 to reverse of: AU221819 from: 1 to: 599

102 glyvaltyrserprohisprotyrgluingluvalseralaleuglyl 118
||||| ::::| ::::| ::::| ::::|
395 GGGATTCATCCACCGACCATCATCAACAACA.....GA 361

118 yllepyrtyrsenglnlletyrglyrrpyryrargvalasnpheglyvali 135
:::|::|::| ::::| ::::| ::::| ::::|
360 TCTACCGTATGGAAAGTCGCTCT.....CACTGAGGAGCA 323

135 leaspluarlglenhisargasnarvglutyrarsapharytyrtyrarg 151
::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
322 AAGAACAGAGGATCCACGAGTTCGATCTTAACCAAGACATCAACCC 273

152 AsnleuAsnIleAlaProAlaGluaspGlyTyrArgLeuAlaGlyPhePr 168
||| ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```



```

80  TyrsertHlyTyrTyle.....TyrValIleal 89
    |||||.....|
67  TACAAACGGCTATTATGTGGAATAAATTCCTAAAGAGTTTTCATCA 116
    |||||.....|
89  aThrAlaProAsnMetPheAsnValAsnAsp..... 99
    |||||.....|
117  TATTCGCGCTGATTTGAAAAMACAGCGTGACTGGAACGACACGCGCA 166
    |||||.....|
100  ..ValIleuGlyValItyserProHisProTyrGluGlnIuValserAla 115
    |||||.....|
167  AACCTCGTCGCGTGTAAACAGACGTCCTACAGACAGGTGTGGCAGTGC 216
    |||||.....|
116  IeuGly.....GlyIleProTyr...SerGlnIleTy 125
    |||||.....|
217  AACGGCAGATACTATTGGCCCGCATGGCACAAAGTTTCTTCATGTGTA 266
    |||||.....|
125  r.....GlyT 127
    |
267  CAACAGATCTATTTCGACGACAGCACACAGCTCAACAGCAGCAGACGT 316
    |||||.....|
127  rPTyArGyAlAsnPhleglyValIleAspGluArgLeuHisArgAsnArg 143
    |||||.....|
317  CGCATCGATTTAAAGCCAGGTAGTCGACAAAGCGCGCGCTCTTTTA 366
    |||||.....|
144  GluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGlu 160
    |||||.....|
367  GAT.....TTCCTCGTTTGA 383
    |||||.....|
160  pGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTyrArgGlu 177
    |||||.....|
384  CGGATATGTTTGGCATCGTCGCCCATACACACACCGCAGTGAATGAG 433
    |||||.....|
177  Lu 177
    ||
434  AA 435

seq_name: gb_est1:AU215966

seq_documentation_block:
LOCUS      AU215966                621 bp      mRNA      linear      EST 17-JUL-2001
DEFINITION AU215966 unpublished oligo-capped cDNA library, stage L1
            Caenorhabditis elegans cDNA clone yK832b05 3', mRNA sequence.
ACCESSION  AU215966
VERSION    AU215966.1 GI:14854123
KEYWORDS   EST.
SOURCE     Caenorhabditis elegans.
ORGANISM   Caenorhabditis elegans.
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
            ; Rhabditidae; Pelodirinae; Caenorhabditis.
REFERENCE  1 (bases 1 to 621)
AUTHORS    Kohara,Y., Shino,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
            and Sugano,S.
TITLE      A complementary view of the C.elegans genome
JOURNAL    Unpublished (2001)
COMMENT    Contact: Yuji Kohara
            Genome Biology Lab.
            National Institute of Genetics
            Yata 111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
            Location/Qualifiers
1..621
   /organism="Caenorhabditis elegans"
   /strain="N2"
   /db_xref="taxon:6239"
   /clone="yK832b05"
   /clone_1fb="unpublished oligo-capped cDNA library, stage
L1"
   /sex="Hermaphrodite"
   /tissue_type="whole animal"
   /dev_stage="L1"

```



```

FEATURES
    source
        Fax: 81-559-81-6855
        Email: tshint@genes.nig.ac.jp.
        Location/Qualifiers
            1..634
                /organism="Caenorhabditis elegans"
                /strain="N2"
                /db_xref="taxon:6239"
                /clone="YK1135f09"
                /elegant="L1 stage"
                /sex="hermaphrodite"
                /tissue_type="whole animal"
                /dev_stage="L1"

BASE COUNT      139 a      141 c      174 g      180 t
ORIGIN
Alignment_scores:
    quality:      80.50      Length:      137
    Ratio:        1.184      Gaps:        7
    Percent Similarity: 49.635      Percent Identity: 24.088

Alignment_block:
    US-09-528-682-1 x BJ138457/rev ..

Align seg 1/1 to reverse of: BJ138457 from: 1 to: 634

102 GYVATYTYSerProHisProTyrGInGluValSerAlaLeuLysG1 118
|||||:|||||:|||||:|||||:|||||:
448 GGGATTCATCCACCGTACCATCATCAACAACA.....CA 414
|||||:|||||:|||||:|||||:|||||:
118 ylleProTyrSerGlnIleTyrGlyTTPYrYrGValaInpHeGlyVal1 135
|||||:|||||:|||||:|||||:|||||:
413 TCTACCGTATTTGGAAAGTCTGCTCT.....CACTGAGGAAGGA 376
|||||:|||||:|||||:|||||:|||||:
135 leAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArg 151
|||||:|||||:|||||:|||||:|||||:
375 AAGAACACAGCGATCCACGAGTTCGATCTTACCACCAAGACCATCACCCC 326
|||||:|||||:|||||:|||||:|||||:
152 AsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheP 168
|||||:|||||:|||||:|||||:|||||:
325 AATGGAGAGATTCGCCAAGTACGG.....TATCGTCAACACAGACTA.. 284
|||||:|||||:|||||:|||||:|||||:
168 oProAspHisGlnAlaTTP.....ArgGluGluProTrrPleHisH 182
|||||:|||||:|||||:|||||:|||||:
283 ...CATCATGCTCTCGTGGAGCCGTTCTCGGACCAAGAAGGCTTTATC 238
|||||:|||||:|||||:|||||:|||||:
182 IAlaPProGlnGlyCysGlyAsnSerSerArgThrIleHrGlyAspTrH 198
|||||:|||||:|||||:|||||:|||||:
237 ACCCTTCGGAGACTGCTCATCACCCA..... 212
|||||:|||||:|||||:|||||:|||||:
199 CysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrG 215
|||||:|||||:|||||:|||||:|||||:
211 ...AACCAAGAGATGGGCCCA.....CGAGAAATATCA 183
|||||:|||||:|||||:|||||:|||||:
215 nserIysValIysArgGlnIlePheSerAspTyrGlnSerGlnValAsp 232
|||||:|||||:|||||:|||||:|||||:
182 ACCCTAAGAGGATCGACACCACTTCACAGACCGGACGAGCTGCTTCA 133
|||||:|||||:|||||:|||||:|||||:
232 leTyrAsnArg 235
|||||:|||||:|||||:|||||:|||||:
132 ACCACACGCCGA 122

seq_name: gb_est1:A0218912

seq_documentation_block:
LOCUS      A0218912      639 bp      mRNA      linear      EST 17-JUL-2001
DEFINITION      A0218912 unpublished oligo-capped cDNA library, stage L1
                Caenorhabditis elegans cDNA clone yk870h06 3', mRNA sequence.
ACCESSION      A0218912
VERSION        A0218912.1      GI:14857069
KEYWORDS
    EST.
    Caenorhabditis elegans.
    SOURCE

```


ORGANISM *Caenorhabditis elegans*

FEATURES

Source

1.639

/organism="Caenorhabditis elegans"

```
/strain="N2"
```

```
/scramble= N2
/db xref="taxon:6239"
```

```
/ad_xlet= lax01:0  
/c]one="vk870h06"
```

```
/clone="yk8/Uu6"
/c)one lib="unpubl)shed o)igo-canned adna libray" atagc
```

```
/clone_lib="unpublis
```

11

```
/sex="Hermaphrodite"
```

BASE COUNT

BASE C
ORIGIN

Category	Count	Percentage
150 c	179	18.3 %

alignment_scores:

Quality: 80.50

Ratio: 1.184

Ratio:	1.104	caps:
Parity:	49.635	Percent Identity:

length: 137

Gaps: 7

Gap: /
Intensity: 24.088

alignment_block:

US-09-528-682-1 x AU218912/rev

Align seg 1/1 to reverse of: AU218912 from: 1 to: 639

[illegible]

OM of: US-09-528-682-1 to: Issued_Patents_NA.* out_format : pfs

Date: Jun 18, 2002 7:28 PM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-Model=framed_p2n.model -Dev=xlp
-O=/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-823-120-5 + 1183.00 2385.79 2.1e-136 711 1 5
-DB=Issued_Patents_NA -OFMT=fasta -SUFFIX=p2n.rni
-GAPOP=12.000 -GAPEXT=4.500 -MINMATCH=0.100 -LOOPCL=0.000
-LOOPEXT=0.000 -GAPEXT=4.500 -GAPEXT=0.050 -XGAPOP=10.000
-XGAPEXT=0.500 -XGAPOP=6.000 -GAPEXT=7.000 -XGAPOP=10.000
-XGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=blsune62 -TRANS=human40.cdi -LIST=1000 -DOALIGN=200
-THR_SCORE=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFM=pfs -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09528682 -CGN1_1_96 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -CGN1_TIMEOUT=30 -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-1
Query Length: 240
Database: Issued_Patents_NA.*
Database sequences: 38353
Search time (sec): 143.210000

Score list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-823-120-5 + 1183.00 2385.79 2.1e-136 711 1 5					
/cgn2_6/ptodata/2/1na/6A.COMB.seq:US-08-823-120-7 + 1088.00 2375.80 1.2e-124 723 1 8					
/cgn2_6/ptodata/2/1na/5A.COMB.seq:US-08-435-605A-1 + 1088.00 2374.99 1.3e-124 777 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-449-045C-1 + 1088.00 2374.94 1.3e-124 774 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-5 + 1073.00 2330.41 1.1e-122 2000 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-7 + 934.00 2038.31 7.6e-106 582 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-9 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-10 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-11 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-12 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-13 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-14 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-15 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-16 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-17 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-18 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-19 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-20 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-21 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-22 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-23 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-24 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-25 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-26 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-27 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-28 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-29 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-30 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-31 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-32 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-33 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-34 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-35 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-36 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-37 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-38 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-39 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-40 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-41 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-42 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-43 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-44 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-45 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-46 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-47 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-48 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-49 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-50 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-51 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-52 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-53 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-54 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-55 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-56 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-57 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-58 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-59 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-60 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-61 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-62 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-63 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-64 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-65 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-66 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-67 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-68 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-69 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-70 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-71 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-72 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-73 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-74 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-75 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-76 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-77 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-78 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-79 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-80 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-81 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-82 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-83 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-84 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-85 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-86 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-87 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-88 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-89 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-90 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-91 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-92 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-93 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-94 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-95 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-96 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-97 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-98 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-99 + 933.00 2036.23 9.9e-106 576 1 1					
/cgn2_6/ptodata/2/1na/5B.COMB.seq:US-08-435-605A-100 + 933.00 2036.23 9.9e-106 576 1 1					

/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-27 + 77.00 136.02 6.90 1 1322	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-28 + 76.50 123.93 32.50 1 3318	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-29 + 76.50 117.13 77.71 1 5864	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-30 + 76.50 117.13 77.71 1 5864	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-31 + 76.50 117.13 77.71 1 5864	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-32 + 76.50 117.13 77.71 1 5864	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-33 + 76.50 117.13 77.71 1 5864	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-34 + 76.50 117.13 77.71 1 5864	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-35 + 76.50 117.13 77.71 1 5864	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-36 + 76.50 117.13 77.71 1 5864	/cgn2_6/ptodata/2/1na/6B.COMB.seq:US-09-823-494-37 + 76.50 11
---	--	--	--	--	--	--	--	--	--	---

/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-739-176-1	70.50	91.42	2.1e+03	165656	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-568-471-2	68.50	70.09	3.2e+04	68750	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-03-036-987A-1	68.50	69.54	3.4e+04	78193
/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-931-869-1	70.50	91.42	2.1e+03	16656	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-03-036-987A-1	68.50	68.26	4.0e+04	8016	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-03-036-987A-1	68.50	68.26	4.0e+04	8016
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-350-399-1	70.50	91.42	2.1e+03	165656	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-03-036-987A-1	68.50	68.26	4.0e+04	8016	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-03-036-987A-1	68.50	68.26	4.0e+04	8016
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-730-379-17	70.50	91.42	2.1e+03	35100	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-730-379-17	70.50	91.42	2.1e+03	35100	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-730-379-17	70.50	91.42	2.1e+03	35100
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-257-669A-17	70.50	82.53	6.6e+03	35100	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-257-669A-17	70.50	82.53	6.6e+03	35100	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-257-669A-17	70.50	82.53	6.6e+03	35100
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-230-371A-17	70.50	79.07	1.0e+04	46899	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-230-371A-17	70.50	79.07	1.0e+04	46899	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-230-371A-17	70.50	79.07	1.0e+04	46899
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-471-115A-7	70.00	120.20	52.42	1362	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-471-115A-7	70.00	120.20	52.42	1362	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-471-115A-7	70.00	120.20	52.42	1362
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-785-396-7	70.00	117.18	77.28	1755	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-785-396-7	70.00	117.18	77.28	1755	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-785-396-7	70.00	117.18	77.28	1755
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-317-785-1	70.00	113.91	117.53	2308	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-317-785-1	70.00	113.91	117.53	2308	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-317-785-1	70.00	113.91	117.53	2308
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-651-086-7	70.00	113.97	120.27	2436	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-651-086-7	70.00	113.97	120.27	2436	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-651-086-7	70.00	113.97	120.27	2436
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-018-760-3	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-018-760-3	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-018-760-3	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-073-384C-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-073-384C-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-073-384C-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-453-359A-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-453-359A-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-453-359A-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-483-043-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-483-043-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-483-043-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-481-238-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-481-238-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-481-238-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-471-066B-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-471-066B-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-471-066B-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-484-956-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-484-956-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-484-956-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-559-491-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-559-491-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-559-491-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-756-386-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-756-386-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-756-386-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-823-516-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-823-516-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-823-516-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-682-853A-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-682-853A-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-682-853A-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-759-038-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-758-314-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-758-314-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-758-314-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-09-350-309-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-09-350-309-2	70.00	112.97	132.49	2496	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-09-350-309-2	70.00	112.97	132.49	2496
/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-07-686-322A-3	70.00	107.90	254.05	3819	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-07-686-322A-3	70.00	107.90	254.05	3819	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-07-686-322A-3	70.00	107.90	254.05	3819
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-002-999-3	70.00	81.42	7.6e+03	35100	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-002-999-3	70.00	81.42	7.6e+03	35100	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-002-999-3	70.00	81.42	7.6e+03	35100
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-770-379-19	70.00	81.42	7.6e+03	35100	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-770-379-19	70.00	81.42	7.6e+03	35100	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-770-379-19	70.00	81.42	7.6e+03	35100
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-737-669A-19	70.00	81.42	7.6e+03	35100	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-737-669A-19	70.00	81.42	7.6e+03	35100	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-737-669A-19	70.00	81.42	7.6e+03	35100
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-330-371A-19	70.00	81.42	7.6e+03	35100	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-330-371A-19	70.00	81.42	7.6e+03	35100	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-330-371A-19	70.00	81.42	7.6e+03	35100
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-676-444-41	69.50	116.88	80.26	1640	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-676-444-41	69.50	116.88	80.26	1640	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-676-444-41	69.50	116.88	80.26	1640
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-077-939-14	69.50	114.98	102.41	1923	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-077-939-14	69.50	114.98	102.41	1923	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-077-939-14	69.50	114.98	102.41	1923
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-461-599-14	69.50	114.98	102.41	1923	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-461-599-14	69.50	114.98	102.41	1923	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-461-599-14	69.50	114.98	102.41	1923
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-461-621-14	69.50	114.98	102.41	1923	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-461-621-14	69.50	114.98	102.41	1923	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-461-621-14	69.50	114.98	102.41	1923
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-465-334-14	69.50	114.87	103.88	1941	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-465-334-14	69.50	114.87	103.88	1941	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-465-334-14	69.50	114.87	103.88	1941
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-625-188-9	69.50	112.66	137.93	2336	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-625-188-9	69.50	112.66	137.93	2336	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-625-188-9	69.50	112.66	137.93	2336
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-228-986-10	69.50	112.05	149.20	2459	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-228-986-10	69.50	112.05	149.20	2459	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-228-986-10	69.50	112.05	149.20	2459
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-101-593-5	69.50	112.05	149.20	2459	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-101-593-5	69.50	112.05	149.20	2459	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-101-593-5	69.50	112.05	149.20	2459
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-465-995A-5	69.50	112.05	149.20	2459	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-465-995A-5	69.50	112.05	149.20	2459	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-465-995A-5	69.50	112.05	149.20	2459
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-465-994C-5	69.50	112.05	149.20	2459	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-465-994C-5	69.50	112.05	149.20	2459	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-08-465-994C-5	69.50	112.05	149.20	2459
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-104-072B-6	69.50	107.41	270.47	3627	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-104-072B-6	69.50	107.41	270.47	3627	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-104-072B-6	69.50	107.41	270.47	3627
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-351-413-7	69.50	107.41	270.47	3627	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-351-413-7	69.50	107.41	270.47	3627	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-351-413-7	69.50	107.41	270.47	3627
/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-025-583-7	69.50	107.41	270.47	3627	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-025-583-7	69.50	107.41	270.47	3627	/cgn2_6/p/oddata/2/1na/5b.COMB	seq:US-09-025-583-7	69.50	107.41	270.47	3627
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-842-936B-47	69.50	107.41	270.59	3628	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-842-936B-47	69.50	107.41	270.59	3628	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-842-936B-47	69.50	107.41	270.59	3628
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-838-973B-47	69.50	107.41	270.59	3628	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-838-973B-47	69.50	107.41	270.59	3628	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-08-838-973B-47	69.50	107.41	270.59	3628
/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-103-820A-2	69.50	118.09	68.73	1351	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-103-820A-2	69.50	118.09	68.73	1351	/cgn2_6/p/oddata/2/1na/6b.COMB	seq:US-09-103-820A-2	69.50	118.09	68.73	1351
/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-261-086-5	69.00	113.52	123.57	1982	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-261-086-5	69.00	113.52	123.57	1982	/cgn2_6/p/oddata/2/1na/5a.COMB	seq:US-08-261-086-5	69.00	113.52	123.57	1982
/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-537-361E-7	69.00	111.34	163.31	2378	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-537-361E-7	69.00	111.34	163.31	2378	/cgn2_6/p/oddata/2/1na/6a.COMB	seq:US-08-537-361E-7				

/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-922-635.3 +	66.50	101.85	553.00	3318	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-922-635.1 +	66.50	101.61	569.15	3385	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-937-931.1 +	66.50	95.54	1.2e+03	5630	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-285-502.1 +	66.50	95.54	1.2e+03	5630	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-385-028.13 +	66.50	95.54	1.2e+03	5630	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-385-028.13 +	66.50	95.54	1.2e+03	5630	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-842-106B.3 +	66.00	83.78	5.6e+03	15079	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-838-973B.3 +	66.00	116.04	89.42	921	1
/csgn2_6/p/odata2/2./1na/PCURS.CONB.seq:PCT-US96-02331.12 +	66.00	111.47	160.75	1351	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-491-772.6 +	66.00	110.65	178.37	1446	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-413-304.8 -	66.00	109.32	211.65	1617	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-991-826.1 -	66.00	108.15	246.01	1784	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-09-143-470.1 -	66.00	108.15	246.01	1784	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-247-475.49 +	66.00	107.52	266.56	1880	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-479-650.49 +	66.00	107.52	266.56	1880	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-191-866D.80 +	66.00	107.52	266.56	1880	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-674-169.49 +	66.00	107.52	266.56	1880	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-985-49B.80 +	66.00	107.43	269.61	1894	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-985-450.7 -	66.00	106.23	314.62	2095	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-261-086.1 +	66.00	101.88	549.82	3017	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-444-792.1 +	66.00	101.88	549.82	3017	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-445-042.1 +	66.00	101.28	593.64	3172	1
/csgn2_6/p/odata2/2./1na/backfilltest.seq:5196511.1 +	66.00	101.28	593.64	3172	1
/csgn2_6/p/odata2/2./1na/PCURS.CONB.seq:PCT-US96-02331.14 +	66.00	97.77	930.65	4255	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-413-304.7 -	66.00	97.64	941.05	4286	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-480-604A.5 +	66.00	90.04	2.5e+03	8133	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-405-196A.5 +	66.00	90.04	2.5e+03	8133	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-915-136.5 +	66.00	90.04	2.5e+03	8133	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-414-926A.6 -	66.00	80.35	8.6e+03	18318	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-926-922.6 -	66.00	80.35	8.6e+03	18318	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-09-253-682.6 -	66.00	80.35	8.6e+03	18318	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-527-657.6 -	66.00	80.35	8.6e+03	18318	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-818-112.17 +	65.50	106.47	305.12	1872	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-818-111.17 +	65.50	106.47	305.12	1872	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-056-556.17 +	65.50	106.47	305.12	1872	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-817-07.7 +	65.50	106.47	305.12	1872	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-817-07.7 +	65.50	103.61	440.06	2378	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-468-558.1 +	65.50	102.70	495.00	2568	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-676-444.1 +	65.50	102.70	495.00	2568	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-804-439A.92 +	65.50	100.62	646.04	3056	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-720-229.92 +	65.50	100.62	646.04	3056	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-804-439A.91 +	65.50	98.62	834.41	3612	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-720-229.91 +	65.50	98.61	880.45	3741	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-541-782.9 -	65.50	92.21	1.9e+03	6176	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-08-911-321.6 +	65.50	92.21	1.9e+03	6176	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-479-637A.1 +	65.50	92.19	1.9e+03	6192	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-083-116.1 +	65.50	92.19	1.9e+03	6192	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-134-116A.1 -	65.50	92.19	1.9e+03	6192	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-09-413-116A.1 -	65.50	92.19	1.9e+03	6192	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-460-751.1 -	65.50	83.42	5.9e+03	12912	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-468-136.4 -	65.50	82.41	6.7e+03	14060	1
/csgn2_6/p/odata2/2./1na/5A.CONB.seq:US-08-445-058A.25 +	65.00	148.60	1.37	50	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-435-605A.37 +	65.00	148.60	1.37	50	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-508-735A.44 -	65.00	120.95	47.60	507	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-201-139.44 -	65.00	120.95	47.60	507	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-129-030.11 +	65.00	119.14	60.54	590	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-020-956.44 -	65.00	114.76	105.36	852	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-030-607.44 -	65.00	114.76	105.36	852	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-439-313.44 -	65.00	114.76	105.36	852	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-967-272.1 -	65.00	110.08	192.01	1261	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-613-182.14 +	65.00	109.92	195.99	1278	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-09-043-123.3 +	65.00	107.56	265.15	1557	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-09-043-123.3 +	65.00	107.56	265.15	1557	1
/csgn2_6/p/odata2/2./1na/PCURS.CONB.seq:PCT-US96-05800.6 +	65.00	107.56	265.15	1557	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-941-445A.1580 -	65.00	106.16	317.63	1752	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-328-501.12 +	65.00	104.93	371.56	1941	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-328-501.12 +	65.00	104.93	371.56	1941	1
/csgn2_6/p/odata2/2./1na/5B.CONB.seq:US-08-637-163B.5 +	65.00	104.50	392.86	2013	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-09-170-354.5 +	65.00	103.99	419.45	2101	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-858-207A.46 +	65.00	103.40	452.28	2207	1
/csgn2_6/p/odata2/2./1na/6B.CONB.seq:US-09-613-182.5 +	65.00	100.37	667.48	2846	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na/6A.CONB.seq:US-08-936-133.5 +	65.00	97.39	977.69	3652	1
/csgn2_6/p/odata2/2./1na					

[illegible]

[illegible]

/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-468-93-43 +	63.00	93.46	1.6e+03	350.05
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-376-84-43 +	63.00	93.46	1.6e+03	350.05
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-185-339-34 +	63.00	93.46	1.6e+03	350.05
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-335-246C-1 +	63.00	92.12	1.9e+03	359.24
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-222-617A-26 +	63.00	79.34	9.8e+03	114.04
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-222-617A-1 +	63.00	78.42	1.1e+04	126.66
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-09-284-878-2 +	63.50	112.23	145.70	663
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-09-201-641-3 +	63.50	107.43	269.56	991
/cgn2_6/p.todata/2./ina./backflist1.seq	:seq:5240849-4 -	62.50	105.51	344.84	1164
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-09-147-926-1 +	62.50	105.36	351.67	1179
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-08-756-299-3 -	62.50	103.72	434.15	1353
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-964-494-3 -	62.50	103.72	434.15	1353
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-889-108-1 +	62.50	101.58	570.88	1618
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-889-108-1 +	62.50	101.58	570.88	1618
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-08-120-601B-1 +	62.50	101.58	570.88	1618
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-08-120-601B-3 +	62.50	101.58	570.88	1618
/cgn2_6/p.todata/2./ina./PCPUS.COMB.seq	:PCR-US94-10358-1 +	62.50	101.58	570.88	1618
/cgn2_6/p.todata/2./ina./PCPUS.COMB.seq	:PCR-US94-10358-3 +	62.50	101.58	570.88	1618
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-188-811-1 -	62.50	100.97	617.97	1704
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-632-740-5 +	62.50	100.02	697.94	1845
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-440-689-5 +	62.50	100.02	697.94	1845
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-122-339-5 +	62.50	100.02	697.94	1845
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-632-740-7 +	62.50	100.00	699.68	1848
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-440-689-7 +	62.50	100.00	699.68	1848
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-09-122-339-7 +	62.50	100.00	699.68	1848
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-08-430-286A-1 +	62.50	98.28	872.70	2135.35
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-202-389-7 +	62.50	98.23	877.71	2143
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-338-685-1 +	62.50	97.51	963.03	2277
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-716-301-3 -	62.50	95.80	1.2e+03	2627
/cgn2_6/p.todata/2./ina./backflist1.seq	:seq:5240849-1 -	62.50	95.26	1.3e+03	2749
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-08-588-576-1 -	62.50	94.70	1.4e+03	2880
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-472-534-1 +	62.50	93.57	1.6e+03	3167
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-965-904B-1 +	62.50	93.25	1.7e+03	3254
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-07-872-634-12 -	62.50	91.43	2.1e+03	3789.83
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-297-494-42 -	62.50	91.43	2.1e+03	3789.83
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-297-494-42 -	62.50	91.43	2.1e+03	3789.83
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-479-532-42 -	62.50	91.43	2.1e+03	3789.83
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-455-525-42 -	62.50	91.43	2.1e+03	3789.83
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-455-525-42 -	62.50	91.43	2.1e+03	3789.83
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-139-491-12 -	62.50	91.43	2.1e+03	3789.83
/cgn2_6/p.todata/2./ina./PCPUS.COMB.seq	:PCR-US92-03222-42 -	62.50	91.43	2.1e+03	3789.83
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-07-872-644-38 -	62.50	90.40	2.4e+03	4133.04
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-297-494-38 -	62.50	90.40	2.4e+03	4133.04
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-297-494-38 -	62.50	90.40	2.4e+03	4133.04
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-479-532-38 -	62.50	90.40	2.4e+03	4133.04
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-455-525-38 -	62.50	90.40	2.4e+03	4133.04
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-139-491-38 -	62.50	90.40	2.4e+03	4133.04
/cgn2_6/p.todata/2./ina./PCPUS.COMB.seq	:PCR-US92-03222-38 -	62.50	90.40	2.4e+03	4133.04
/cgn2_6/p.todata/2./ina./PCPUS.COMB.seq	:PCR-US92-03222-38 -	62.50	90.40	2.4e+03	4133.04
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-417-089-5 +	62.50	83.33	5.9e+03	7470
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-695-651-5 +	62.50	83.33	5.9e+03	7470
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-930-285-5 +	62.50	83.33	5.9e+03	7470
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-08-695-421-5 +	62.50	83.33	5.9e+03	7470
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-08-695-421-5 +	62.50	83.33	5.9e+03	7470
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-09-073-492-1 +	62.50	74.65	1.8e+04	1546.65
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-08-846-762-1 +	62.50	69.19	3.6e+04	2441.64
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-846-762-1 +	62.50	69.19	3.6e+04	2441.64
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-09-171-461-1 +	62.50	62.22	8.4e+04	4380.44
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-079-913-1 -	62.50	56.13	7.9e+05	7292.92
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-447-810-35 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-487-933-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-487-933-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-480-810-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-508-735-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-848-251-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-848-251-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./5B.CONB.seq	:US-08-466-047-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-120-130-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-120-130-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-115-252-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-966-515-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-08-966-515-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-120-128-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-120-128-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6A.CONB.seq	:US-09-120-139-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-09-120-131-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./6B.CONB.seq	:US-09-120-131-25 -	62.00	117.31	75.98	395
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-181-271A-7 +	62.00	105.99	324.59	1020
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-449-813-7 +	62.00	105.99	324.59	1020
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-449-813-7 +	62.00	105.99	324.59	1020
/cgn2_6/p.todata/2./ina./5A.CONB.seq	:US-08-444-803-7 +	62.00	105.99	324.59	1020

cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-449-043-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-08-444-818-88 +	62.00	80.66	8.0e+03	8316
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-456-265A-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-08-827-308-1 +	62.00	80.66	8.3e+03	8517
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-455-416-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-500-358-1 +	62.00	80.66	8.3e+03	8517
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-455-244-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-458-809-1 +	62.00	80.66	8.3e+03	8517
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-457-364-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-08-444-818-137 +	62.00	80.02	9.0e+03	8987
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-456-262-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-08-444-818-122 +	62.00	79.76	9.3e+03	9188
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-456-240-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-08-444-818-123 +	62.00	79.51	9.6e+03	9379
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-455-736-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-08-444-818-176 +	62.00	79.51	9.6e+03	9379
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-911-217-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-388-874-1 +	62.00	79.48	9.7e+03	9401
cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-550-600-7 +	62.00	105.99	324.59	1020	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-07-910-760-9 +	62.00	79.48	9.7e+03	9401
cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-08-555-568B-18 +	62.00	103.30	457.84	1277	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-440-519-9 +	62.00	79.48	9.7e+03	9401
cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-09-519-223-18 +	62.00	102.94	479.84	1277	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-08-440-545-9 +	62.00	79.00	1.0e+04	9785
cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-08-886-886-10 +	62.00	101.12	609.18	1534	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-319-387-1 -	62.00	79.00	1.0e+04	9785
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-850-880-1 -	62.00	101.12	609.18	1534	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-338-907-73 +	62.00	73.31	2.1e+04	15766
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-944-916-1 -	62.00	101.12	608.18	1534	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-218-207-73 +	62.00	73.31	2.1e+04	15766
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-814-877-1 -	62.00	101.12	608.18	1534	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-097-889-2 +	62.00	72.72	2.3e+04	15659
cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-09-146-283-1 +	62.00	100.70	639.15	1588	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-377-885-1 +	62.00	72.72	2.3e+04	15659
cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-08-579-823A-1 +	62.00	100.70	639.15	1588	1	cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-850-880-11 +	62.00	70.97	2.8e+04	19188
cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-344-195-1 +	62.00	100.70	639.15	1588	1	cgcn2_6/p.todata/2/1na/5B.COMB.seq:US-08-944-916-11 +	62.00	70.97	2.8e+04	19188
cgcn2_6/p.todata/2/1na/6A.COMB.seq:US-09-033-055A-1 +	62.00	99.74	773.57	1722	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-08-976-259-14 -	62.00	68.97	3.6e+04	22677
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-021-608D-7 +	62.00	99.19	776.27	1803	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-306-691B-19 -	62.00	63.76	7.0e+04	35131
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-126-160-7 +	62.00	99.19	776.27	1803	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-306-691B-19 -	62.00	63.76	7.0e+04	35131
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-333-358-1 -	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-338-907-183 +	62.00	62.83	7.8e+04	37993
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/6B.COMB.seq:US-09-218-207-183 +	62.00	62.83	7.8e+04	37993
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-477-728-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-477-728-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-474-940-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-474-940-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61.50	114.49	109.06	456
cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-634-694-1 +	62.00	98.24	876.59	1932	1	cgcn2_6/p.todata/2/1na/5A.COMB.seq:US-08-484-537-100 +	61			

[illegible]

/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-508-004-2 +	60.50	80.37	8.6e+03	6611	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-481-130-27 +	60.00	88.99	2.9e+03	2927
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-402-068-2 +	60.50	80.37	8.6e+03	6611	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-656-964A-27 +	60.00	88.99	2.9e+03	2927
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-402-068-2 +	60.50	80.37	8.6e+03	6611	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-485-604A-27 +	60.00	88.99	2.9e+03	2927
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-987-439-2 +	60.50	80.31	8.6e+03	6702	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-27 +	60.00	88.99	2.9e+03	2927
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-570-311-9 +	60.50	79.87	9.2e+03	6895	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-261-206A-71 +	60.00	87.54	3.5e+03	3306
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-353-485-9 +	60.50	79.87	9.2e+03	6895	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-347-468-1 +	60.00	87.54	3.5e+03	3332
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-718-388-6 +	60.50	78.36	1.1e+04	7824	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-18 +	60.00	86.73	3.8e+03	3357
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-076-011-1 +	60.50	69.07	3.6e+04	17041	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-691-814-38 +	60.00	85.73	4.4e+03	3846
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-07-642-734C-3 -	60.50	67.02	4.7e+04	20235	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-23 +	60.00	85.18	4.7e+03	3855
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-439-009A-3 -	60.50	62.93	5.7e+04	15331	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-392-459-21 +	60.00	84.22	5.3e+03	4365
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-128-155-16 +	60.50	41.18	6.3e+05	176373	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-312-288B-9 -	60.00	85.18	4.7e+03	4029
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-470-179-195 +	60.00	111.96	150.83	427	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-188-930-223 +	60.00	108.94	232.21	550	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-846-021A-4 -	60.00	108.64	230.93	564	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-998-416-730 -	60.00	107.23	276.88	635	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-935-263-1 +	60.00	105.88	329.19	711	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-245-808-7 +	60.00	104.54	390.55	795	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-287-442-4 -	60.00	102.01	540.48	983	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-459-701-4 -	60.00	102.01	540.48	983	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-460-298-4 -	60.00	102.01	540.48	983	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-459-174-4 -	60.00	102.01	540.48	983	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-811-949-64 +	60.00	99.93	705.58	1170	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-261-206A-75 +	60.00	99.13	781.71	1251	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-811-949-48 +	60.00	98.55	842.76	1314	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-811-949-62 +	60.00	97.63	948.00	1419	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-753-007A-7 +	60.00	97.16	1.0e+03	1476	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-398-486-7 +	60.00	97.16	1.0e+03	1476	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-659-166-1 -	60.00	96.32	1.1e+03	1584	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-811-949-64 +	60.00	96.16	1.1e+03	1605	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-07-726-607C-2 +	60.00	96.06	1.2e+03	1618	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-07-843-949A-2 +	60.00	96.06	1.2e+03	1618	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-218-978-2 +	60.00	96.05	1.2e+03	1620	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-977-847-2 -	60.00	96.05	1.2e+03	1620	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-195-021-2 -	60.00	95.74	1.2e+03	1662	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-807-342B-3 -	60.00	95.29	1.3e+03	1727	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-811-949-64 +	60.00	95.21	1.3e+03	1738	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-811-949-64 +	60.00	94.48	1.4e+03	1848	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-376-856-1 +	60.00	94.02	1.5e+03	1920	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-136-574A-46 +	60.00	93.81	1.5e+03	1955	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-883-795A-39 +	60.00	93.36	1.6e+03	2029	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-811-949-42 +	60.00	92.95	1.7e+03	2101	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-119-512-3 +	60.00	92.60	1.8e+03	2162	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-488-015B-3 +	60.00	92.60	1.8e+03	2162	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-974-549A-3 +	60.00	92.53	1.8e+03	2176	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-07-726-607C-1 +	60.00	92.25	1.9e+03	2228	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-07-843-949A-1 +	60.00	92.25	1.9e+03	2228	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-718-978-1 +	60.00	92.25	1.9e+03	2228	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-753-007A-31 +	60.00	92.03	1.9e+03	2268	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-398-486-31 +	60.00	92.03	1.9e+03	2268	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-767-026-6 +	60.00	91.53	2.1e+03	2366	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-245-808-2 +	60.00	91.27	2.1e+03	2418	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-753-007A-1 +	60.00	91.08	2.2e+03	2457	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-09-398-486-1 +	60.00	91.03	2.2e+03	2467	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-920-812-14 -	60.00	91.01	2.2e+03	2471	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-920-838-14 -	60.00	91.01	2.2e+03	2471	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-362-577C-14 -	60.00	91.01	2.2e+03	2471	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-920-828-14 -	60.00	91.01	2.2e+03	2471	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-920-828-14 -	60.00	91.01	2.2e+03	2471	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-724-194-2 +	60.00	90.89	2.3e+03	2497	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-481-130-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-636-984A-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-485-604-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+03	6090
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-612-351-3 -	60.00	80.14	8.9e+03	6143
/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-487-555-25 +	60.00	89.63	2.6e+03	2775	1	/cgn2_6/pdata/2/1na/5A.COMB.seq:US-08-566-338-35 +	60.00	80.24	8.9e+0	


```

/cgn2_6/ptodata/2/lna/5A.COMB.seq:US-08-446-794A-1 + 59.50 87.43 3.5e+03 3040 1
/cgn2_6/ptodata/2/lna/5A.COMB.seq:US-08-750-007-2 + 59.50 87.43 3.5e+03 3040 1
/cgn2_6/ptodata/2/lna/5B.COMB.seq:US-08-945-024-1 + 59.50 87.43 3.5e+03 3040 1
/cgn2_6/ptodata/2/lna/Backfiles1.seq:5244792-1 + 59.50 85.85 4.3e+03 3472 1
/cgn2_6/ptodata/2/lna/5B.COMB.seq:US-08-630-915A-33 + 59.50 83.89 5.5e+03 4091 1
/cgn2_6/ptodata/2/lna/5B.COMB.seq:US-09-586-719-11 - 59.50 83.58 5.7e+03 4198 1
/cgn2_6/ptodata/2/lna/5A.COMB.seq:US-08-004-139B-2 + 59.50 82.21 6.8e+03 4707 1
/cgn2_6/ptodata/2/lna/5B.COMB.seq:US-08-811-492-2 + 59.50 82.21 6.8e+03 4707 1
/cgn2_6/ptodata/2/lna/PCTUS.COMB.seq:PCT-US96-10545A-2 + 59.50 82.21 6.8e+03 4707 1
/cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-436-332B-9 + 59.50 81.84 7.2e+03 4858 1
/cgn2_6/ptodata/2/lna/6A.COMB.seq:US-09-073-354-6 + 59.50 80.71 8.3e+03 5339 1

```

```

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-823-120-5

```

```

seq documentation block:
; Sequence 5, Application US/08823120
; Patent No. 6149919
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizzia, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,120
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,003
; FILING DATE: 11-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..711
; US-08-823-120-5

```

alignment_scores:

```

Quality: 1183.00 Length: 240
Ratio: 5.189 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 93.333

```

alignment block:

```

US-09-528-682-1 x US-08-823-120-5 ..
Align seg 1/1 to: US-08-823-120-5 from: 1 to: 711

```

```

seq documentation block:
; Sequence 7, Application US/08823120
; Patent No. 6149919
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizzia, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines

```



```

1 AsnGlySpArgLeuTyTyrArgAlaSerLeuArgProPheAspGluLeu 17
1  : : : : : : : : : : : : : : : : : : : : : : : : : :
1  AATGATGATAGATTATATCGGCGAGATTCTGAACTCCGTGATGAATTA 50
17 satgSeGclgLyLeuMetProAlagGlyHisAsnGluTyrPheAspArg 34
17 : : : : : : : : : : : : : : : : : : : : : : : : : :
51 GCATCTGAGGTGGCTTTATGCGAAGGGAGGAGTAGTACTTTGACCGAG 100
51 : : : : : : : : : : : : : : : : : : : : : : : : : :
34 LyrHrgImeAsnIleAsnLeuTyTAspHisAlaArgLyrHrgIlnTr 50
34 : : : : : : : : : : : : : : : : : : : : : : : : : :
101 GTACTCAATGAATGATTCACCTTTATGATCATGCAAGAGGAATCGACG 150
101 : : : : : : : : : : : : : : : : : : : : : : : : : :
51 GlyPheValArgTyrAspAspGlyTyrValSerTrnSerLeuSerLeuAl 67
51 : : : : : : : : : : : : : : : : : : : : : : : : : :
151 GGATTCTTGAGCGACGATGATGATATCTTCCACTCAATTAGTTTGAG 200
151 : : : : : : : : : : : : : : : : : : : : : : : : : :
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerTrnTyr 84
67 : : : : : : : : : : : : : : : : : : : : : : : : : :
201 AAGTCCCACTTACTGGGTCAACATATATTCCTGTCATTCCTATCAT 250
201 : : : : : : : : : : : : : : : : : : : : : : : : : :
84 yTlleTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
84 : : : : : : : : : : : : : : : : : : : : : : : : : :
251 ATATTATATGTTATAGCCATCGACCAACATGTTTAACTGTATATGATGA 300
251 : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

1 seq_documentation_block:
2   Sequence 1, Application US/08435605A
3   Patent No. 5874287
4
5   GENERAL INFORMATION:
6     APPLICANT: Burnette, W. Neal
7     APPLICANT: Kaslow, Harvey R.
8     TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
9     TITLE OF INVENTION: SUBUNIT ANALOGS
10    NUMBER OF SEQUENCES: 57
11
12    CORRESPONDENCE ADDRESSES:
13      ADDRESSEE: Amgen Inc.
14      STREET: 1840 De Havilland Drive
15      CITY: Thousand Oaks
16      STATE: California
17
18    COUNTRY: USA
19    ZIP: 91320-1789
20
21    COMPUTER READABLE FORM:
22      MEDIUM TYPE: Floppy disk
23      COMPUTER: IBM PC compatible
24      OPERATING SYSTEM: PC-DOS/MS-DOS
25
26    SOFTWARE: Patentin Release #1.0, Version #1.30
27
28    CURRENT APPLICATION DATA:
29      APPLICATION NUMBER: US/08/435,605A
30      FILING DATE: 05-MAY-1995
31      CLASSIFICATION: 435
32
33    ATTORNEY/AGENT INFORMATION:
34      NAME: Mazza, Richard J., 657
35      REGISTRATION NUMBER: 27,657
36      REFERENCE/DOCKET NUMBER: A-196B
37
38    INFORMATION FOR SEQ ID NO: 1:
39      SEQUENCE CHARACTERISTICS:
40        LENGTH: 774 base pairs
41        TYPE: nucleic acid
42        STRANDEDNESS: single

```


TOPLOGY: linear
MOLECULE TYPE: CDNA
US-08-435-605A-1

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x US-08-435-605A-1 ..

Align seg 1/1 to: US-08-435-605A-1 from: 1 to: 774

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleu 17
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
55 AATGATGATAGATTATATCGGCAGATTCTAGACCTCTGATGAATAATA 104
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  ::::::::::::::::::::
105 GCAGTCAGGCGTCTTATGCCAAGACAGACAGAGTACTTGACCGAG 154
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
155 GTACTCAAAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGAC 204
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
205 GGATTTGTTGGCAGCATGATGATATGTTCCACCTCATTTAGTTGAG 254
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
255 AAGTCCCACTAGTGGTCAACTATATGTCGTGCTGATTCATCTATAT 304
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
84 YrLLeuTyrValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
305 ATATATATGTTATATAGCCACTGACCCCAACATGTTTAACTTAATGAT 354
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
355 TTAGGGGCGATACAGTCTCTCATCGAGATGAACAAGAACTTCTGCTTAG 404
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
405 TGGGATTTCCATCTCCCAATATATATGATGATGATCGATTCATTTGGG 454
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
455 TGGTGTATGACAAATTACATCGTAATAGGGGCTACAGAGATGATATAC 504
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
505 AGTAACTTATGATATGCTCCAGCAGCAGATGTTATGATGTCGACAGTT 554
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
167 ePrProAspHisGlnAlaThrArgGluGluProTyrIleHisAlaPhe 184
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
555 CCTCCGAGCATATAGAGCTTGAGGAGAGAGCCGTGATTCATCATGAC 604
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
605 CGCGGGGTGTGGAAATGCTCCAAGATCATGATGATTAATCTTGGAT 654
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
655 GAAAAAACCCAAAGTCTAGTGTAAATTCCTTGACGAATACCAATCTAA 704
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
705 AGTTAAAGCAATATATTTTCAGGCTATCATCTGATATTGATACACATA 754
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
234 snArgIleArgAspGluLeu 240
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||

```

755 ATAGAAATTAGATGAATT 774

seq_name: /cgn2_6/Plodata/2/lna/5A_COMB.seq:US-08-449-045C-1

seq_documentation_block:

Sequence 1, Application US/08449045C
Patent No. 5770203

GENERAL INFORMATION:

APPLICANT: Burnette, Neal W.

APPLICANT: Kaslow, Harvey R.

TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESS: Amgen Inc.

STREET: 1840 De Havilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/449, 045C

FILING DATE: 24-MAY-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/694,733

FILING DATE: 02-MAY-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/271,222

FILING DATE: 06-JUL-1994

ATTORNEY/AGENT INFORMATION:

NAME: Mazza, Richard J.

REGISTRATION NUMBER: 27,657

REFERENCE/DOCKET NUMBER: A-196C

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 777 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..777

US-08-449-045C-1

alignment_scores:

Quality: 1088.00 Length: 240

Ratio: 4.814 Gaps: 0

Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:

US-09-528-682-1 x US-08-449-045C-1 ..

Align seg 1/1 to: US-08-449-045C-1 from: 1 to: 777

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleu 17
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
55 AATGATGATAGATTATATCGGCAGATTCTAGACCTCTGATGAATAATA 104
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  ::::::::::::::::::::
105 GCAGTCAGGCGTCTTATGCCAAGACAGACAGAGTACTTGACCGAG 154
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
155 GTACTCAAAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGAC 204
  ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||

```



```
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||.....|
205 GGATTGTTAGGCACATGATGATGTTTCCACCTCAATTAGTTTGG 254
67 gserAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
|||||.....|
255 AAGTCCCACTTAGTGTCAAACTATATGCTGTCATCTTACTTATT 304
84 yrlleTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||.....|
305 AATATATGTATTAGCCACTGCACCCACATGTTTAACGTTAATGATGTA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnValSerAlaLeuG 117
|||||.....|
355 TTAGGGGCATACAGTCTCATCCAGATGAACAGAAAGTTTCTGCTTAA 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhel 134
|||||.....|
405 TGGGATTCATACCTCCCAATATATGATGATGATTCAGTTTGGGG 454
134 allleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||.....|
455 TGGTGTATGACAAATTCATCGTAATAGGGGCTACAGATAGATATTA 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAla 167
|||||.....|
505 AGTAACTTGAATATGCTCCACGACGACAGATGTTATGATGGCAGG 554
167 eProProAsnHisGlnAlaThrPargGluGluProTyrIleHisAla 184
|||||.....|
555 CCCTCGGAGCATAGAGCTTGGAGGAGAGCCGTGATTCATCATGCG 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCys 200
|||.....|
605 CGCCGGGTGTGGGAATGCTCCAAAGATCATCGATGATACTTGCAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGln 217
|||||.....|
655 GAAATAACCCAAAGTCTAGGTGTAATAATTCCTGACGAATACCAAT 704
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIle 234
|||||.....|
705 AGTTAAAGCAAAATATTTTCAAGCTATCATCTGATGATGATACACA 754
234 snArgIleArgAspGluLeu 240
|||||.....|
755 ATGAATTAAGATGAATTA 774
seq_name: /cgn2_6/ptodata/2/lna/backfile1.seq:5223610-1
seq_documentation_block:
: Patent No. 5223610
: APPLICANT: Burton, Frank H.; Sutcliffe, Gregor
: TITLE OF INVENTION: CHOLERA TOXIN GENE REGULATED BY GROWTH
: HORMONE PROMOTER
: NUMBER OF SEQUENCES: 18
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/528, 852
: FILING DATE: 18-MAY-1990
: SEQ ID NO: 1
: LENGTH: 2020
5223610-1
```

```
alignment_scores:
  Quality: 1073.00      Length: 240
  Ratio: 4.790          Gaps: 0
  Percent Similarity: 93.333  Percent Identity: 80.833
alignment_block:
US-09-528-682-1 x 5223610-1
```

```
Align seg 1/1 to: 5223610-1 from: 1 to: 2020
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspIleTyr 17
|||.....|
570 AATGATGATTAAGTTAATCGGCGAGATTCAGACCTCTGATGATAAATA 619
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||.....|
620 GCAGTCAGGTGTCTATGCGCAAGAGACAGAGTACTTTGACCCGAG 669
34 IyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGln 50
|||||.....|
670 GTACTCAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGAG 719
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||.....|
720 GGATTTGTTAGCGCAGATGATGATATGTTTCCACCTCAATGATGTT 769
67 gserAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
|||||.....|
770 AAGTCCCACTTAGTGTGCTCAAACTATATGCTGTCATCTTACTTATT 819
84 yrlleTyValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||.....|
820 AATATATGTATTAGCCACTGCACCCACATGTTTAACGTTAATGATGTA 869
101 LeuGlyValTyrSerProHisProTyrGluGlnValSerAlaLeuG 117
|||||.....|
870 TTAGGGGCATACAGTCTCATCCAGATGAACAGAAAGTTTCTGCTTAA 919
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPhel 134
|||||.....|
920 TGGGATTCATACCTCCCAATATATGATGATGATTCAGTTTGGGG 969
134 allleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||.....|
970 TGGTGTATGACAAATTCATCGTAATAGGGGCTACAGATAGATATTA 1019
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAla 167
|||||.....|
1020 AGTAACTTGAATATGCTCCACGACGACAGATGTTATGATGGCAGG 1069
167 eProProAsnHisGlnAlaThrPargGluGluProTyrIleHisAla 184
|||||.....|
1070 CCCTCGGAGCATAGAGCTTGGAGGAGAGCCGTGATTCATCATGCG 1119
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCys 200
|||.....|
1120 CGCCGGGTGTGGGAATGCTCCAAAGATCATCGATGATACTTGCAT 1169
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGln 217
|||||.....|
1170 GAAATAACCCAAAGTCTAGGTGTAATAATTCCTGACGAATACCAAT 1219
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIle 234
|||||.....|
1220 AGTTAAAGCAAAATATTTTCAAGCTATCATCTGATGATGATACACA 1269
234 snArgIleArgAspGluLeu 240
|||||.....|
1270 ATGAATTAAGATGAATTA 1289
seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-5
seq_documentation_block:
: Sequence 5, Application US/08435605A
: Patent No. 5874287
: GENERAL INFORMATION:
: APPLICANT: Burnette, W. Neal
: APPLICANT: Kaslow, Harvey R.
: TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
: SUBUNIT ANALOGS
: NUMBER OF SEQUENCES: 57
```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,605A
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196B
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 582 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-435-605A-5

```

```

alignment_scores:
Quality: 934.00 Length: 193
Ratio: 5.049 Gaps: 0
Percent Similarity: 95.855 Percent Identity: 86.528

```

```

alignment_block:
US-09-528-682-1 x US-08-435-605A-5 ..
Align seg 1/1 to: US-08-435-605A-5 from: 1 to: 582

```

```

1 Asnglyaspargleutyrrarglaaspserargproproaspclutlely 17
|||||
1 AATGATGATTAAGTATATATCGGCAAGATTTCTAGACCTCGATGAATAAA 50
17 sarqserglyglyleumetproargglyhisnsglutyrrpheasparg 34
|::|
51 GCAGTCAGGTGCTTATGCCAAGAGACAGTGTGCTTGGACCGAG 100
34 lythglmetasnleasnleutyrrasphsialaargglythrlnthr 50
|||||
101 GTACTCAATGATATCAACCTTTATGATCATGCAAGAGAACTCAGACG 150
51 glyphevalargtyrraspspalytyrrvalserthrserleuuar 67
|||||
151 GGATTTGTTAGGACAGATGATGATTTCCACCTCAATTAGTTTGAG 200
67 gseralanhlsleualaglylnserlleuuserglytyrserthrtyr 84
|||||
201 AAGTCCCACTTAGTGGTCACAACATATATGCTGCTCATTTACTTATT 250
84 yrlletryvalllealathralaproasmetpheasnvalasnspval 100
|||||
251 AATATATGTTATAGCCACGACCAACATGTTTAACGTTAATGATGTA 300
101 leuglyvaltyrserprohisprotyrlynglunvalseralaleucl 117
|||||
301 TTAGGGGATACAGTCTTCATCCAGATGAACAAGATTTCGCTTTAGG 350
117 yglylleprotyrserglnletryglytyrptyrrargvalasnphgly 134
|||||
351 TGGGATTCATCTCCCAATATATGATGATGATGATGATGATGATGATG 400
134 allleaspluarleuuhisarqasnargglutyrrargspargtyrtyr 150

```

```

|||||
401 TCCTTGATGAACAATTAATCATCTAATAGGGCTACAGAGATGATATAC 450
151 Argasnleuasnllealaproalaglunbspglytyrrargleualaglyph 167
|||||
451 AGTAACCTTAGATATGCTCCACAGACAGATGCTTATGATTTGGCAGGTT 500
167 eProProasphslnalalatrparglunluprothrphleishsialap 184
|||||
501 CCTCCGGAGCATAGAGCTTGAGGGAAGCCGTTGATTCATCATGCAC 550
184 roglnglycsglylnserserargthr 193
|||||
551 CCGCGGTTGTGGGAATGCTCCAGATCA 579

```

```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-435-605A-7

```

```

seq_documentation_block:
Sequence 7, Application US/08435605A
Patent No. 5874287
GENERAL INFORMATION:
APPLICANT: Burnette, W. Neal
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
NUMBER OF INVENTION: SUBUNIT ANALOGS
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,605A
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196B
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-435-605A-7

```

```

alignment_scores:
Quality: 933.00 Length: 192
Ratio: 5.071 Gaps: 0
Percent Similarity: 95.833 Percent Identity: 86.979

```

```

alignment_block:
US-09-528-682-1 x US-08-435-605A-7 ..

```

```

Align seg 1/1 to: US-08-435-605A-7 from: 1 to: 576

```

```

1 Asnglyaspargleutyrrarglaaspserargproproaspclutlely 17
|||||
1 AATGATGATTAAGTATATCGGCAAGATTTCTAGACCTCGATGAATAAA 50
17 sarqserglyglyleumetproargglyhisnsglutyrrpheasparg 34
|::|
51 GCAGTCAGGTGCTTATGCCAAGAGACAGTGTGCTTGGACCGAG 100

```



```

34  LyrhgrlnmetasnlleasleuTYrAspHisAlaArglyThrglnThr 50
    |||||||
101  GTACTGAAATGAATATCAACCTTATGATCATCAAGAGACCTCAGACG 150
    |||||||
51  GlyPheValArgTYrAspArglyTYrValSerThrSerLeuSerLeuAr 67
    |||||||
151  GGATTTGTTGGCAGCATGATGATGTTCCACCCCAATTAAGTTTGAG 200
    |||||||
67  gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYr 84
    |||||||
201  AAGTGCACCATGTAGTGGTCAACATATATGTGTCGTCATTCATTTAT 250
    |||||||
84  yrlleTYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
    |||||||
251  ATATATATGTTATAGCCATGTCACCCCAACATGTTTAAAGTTAATGA 300
    |||||||
101  LeuGlyValTYrSerProHisProTYrGlnGlnIleValSerAlaLeuG 117
    |||||||
301  TTAGGGGCATACAGTCTCATCCAGATGACAGAAAGACTTCTGCTTAGG 350
    |||||||
117  yGlyIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPheGly 134
    |||||||
351  TGGGATTCCTACTCCCAATATATGATGATGATGATGATGATGATGATG 400
    |||||||
134  alileasprguargleuHisArgAsnArgGlyTYrArgAspArgTYrTYr 150
    |||||||
401  TGCTTGATGACAAATATACATGATATAGGGGATGATGATGATGATGAT 450
    |||||||
151  ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGly 167
    |||||||
451  AGTACTTATATATGCTCCAGCAGCAGATGTTATGATGATGATGATGAT 500
    |||||||
167  eProProAspHisGlnAlaThrArgGlnIleProTYrIleHisHisAla 184
    |||||||
501  CCTCCGAGCATAGAGCTTGAGGAGAAAGCGGTGATTCATCATCATGAC 550
    |||||||
184  roGlnGlyCysGlyAsnSerSerArg 192
    |||||||
551  CGCCGGGTGTGGAAATGCTCCAGA 576
    |||||||
seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-9

seq_documentation_block:
; Sequence 9, Application US/08435605A
; Patent No. 5874287
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
; TITLE OF INVENTION: SUBUNIT ANALOGS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,605A
; FILING DATE: 05-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mazza, Richard J.
; REGISTRATION NUMBER: 27,657
; REFERENCE/DOCKET NUMBER: A-1968
; INFORMATION FOR SEQ ID NO: 9:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-435-605A-9

alignment_scores:
  Quality: 157.00      Length: 44
  Ratio: 3.925        Gaps: 0
  Percent Similarity: 90.909  Percent Identity: 65.909

alignment_block:
US-09-528-682-1 x US-08-435-605A-9  ..

Align seg 1/1 to: US-08-435-605A-9 from: 1 to: 138

197 AspThrCysAsnGlnGluThrglnAsnLeuSerThrIleTYrLeuArgG 213
    |||||||
7  AATACCTGCGATGAAAAAACCAAGCTAGGTAAATAATCCCTTGACGA 56
    |||||||
213 uTYrGlnSerIleValysArgGlnIlePheSerAspTYrGlnSerGlu 230
    |||||||
57 ATACCAATCTAAACTTAAAGACAAATATTTTCAGGCTATCAATCTGATA 106
    |||||||
230 alAspIleTYrAsnArgIleArgAspGluLeu 240
    |||||||
107 TTGATACACATAATATAGATTAAGATGATGATA 138
    |||||||
seq_name: /cgn2_6/ptodata/2/lna/backfiles1.seq:5223610-4

seq_documentation_block:
; Patent No. 5223610
; APPLICANT: Burton, Frank H.; Sutcliffe, Gregor
; TITLE OF INVENTION: CHOLERA TOXIN GENE REGULATED BY GROWTH
; HORMONE PROMOTER
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/528,852
; FILING DATE: 18-MAY-1990
; SEQ ID NO: 4
; LENGTH: 4207
5223610-4

alignment_scores:
  Quality: 90.50      Length: 338
  Ratio: 0.730        Gaps: 16
  Percent Similarity: 36.686  Percent Identity: 20.710

alignment_block:
US-09-528-682-1 x 5223610-4  ..

Align seg 1/1 to: 5223610-4 from: 1 to: 4147

5  LeuTYrArgAlaAspSerArgProProAspGlnIleLysArgSerGlyG 21
    |||||||
627 GTATACCGGTATGATCCCGCCCGCGGAGAGACTTTCCAGAAC...GG 673
    |||||||
21  yLeuMetProArgGlyHisAsnGlnTYrPheAspArgGlyThrglnMeta 38
    |||||||
674 ATTCACGCGCTGGGAAACAACGAC..... 698
    |||||||
38  snIleAsnLeuTYrAspHisAlaArgGlyThrglnThrglyPheValArg 54
    |||||||
699  ....AATGTCGACCATCTGACCGGAGAGCTTCTCCAGTGGCAGC 743
    |||||||
55  TYrAspAspGlyTYrValSerThrSerLeuSerLeuArg..... 67
    |||||||
744 AGCAACAGCGCTTTCGTCTCCACGACGAGCGCGGTATACGAGGT 793
    |||||||

```



```

68 .....SerAlaHisLeuAlaGly 74
794 CTATCTGAAACATGCATGCAGAAAGCGGTCGAGCGGCGGCGCGGCA 843
74 InsertIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIle..... 88
844 GGGGC.....ACGGCCACTTTCATCGGCTACATCATGGAAGTCCGGGCC 887
89 .....AlaThrAlaProAsnMetPheAsnValAsnAspVal 100
888 GACAACAATTTCTACGGGGCGCGGAGCTCTACTTTCGAATACGTGCACAC 937
100 lLeuGlyValTyrSerProHis.....ProTyrG 110
938 TTATGGCGACAATGCCGCGCTATCTCGCGCGCGCGCTGCGCACTAAC 987
110 lGlnGlnValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGly 126
988 AGACGAAATATCTGGCACACCGCGCATTCGCGCGGAAACATCCGACAG 1037
127 TrpTyrArgVal..... 130
1038 GTAACGGGCGCTATCAACAGGCATCAGCGCGGAGAACACAGACACGGA 1087
131 .....AsnPheGlyValIleAspGluArgLeuHisArgAsn...ArgG 144
1088 CTATTCGACGCTCGCTACGTACAGCAGACATCTCGCGCAATCCCAAC 1137
144 lLysTyrArgAspArg..... 148
1138 CTTACACATTCGCGAAGTCCGTACGCTGATCTCGGCACATTGGTCAT 1187
148 ..... 148
1188 GGGCCCGGTATAGCGCTTGCATGGCGCGGCGGCGGAAAGCTCCGAGGC 1237
149 .....TyrT 150
1238 CATGGACGCTGTCTCGAACCGCGCGGAGCGATGTCTCTCGTACT 1287
150 YR..... 150
1288 ACGAAACGATCGCTATTCTGTTCAGACCTGGCCAGCCCGCCCAACTC 1337
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly.. 166
1338 CGGTAATTGAACACATGCGGATGCGGAGCAAGACGTCGCGCCAGCAT 1387
167 ....PheProAspHisGlnAlaTTPArgGluProTPrIleHisH 182
1388 CGTCATTCGCGCGAGAACAGATTACCCAG.....CATG 1422
182 lAlaIleProGlnGly...CysGlnAsnSerSerArgThrIleThr..... 195
1423 GCAAGCCCTATGACGCTGCGGCAACAGACCTGCGCTGACCGGTGGCG 1472
196 .....GlyAspThrCysAsnGluIleThrGlnAsnLeuSe 207
1473 GAATTGGCGCGGCGAGCGGAT..... 1493
207 rThrIleTyrLeuArgGluTyrGlnSerLysValLysArg.....GlnI 222
1494 .....CTGCAGAGTACACTGCTCATGTGACGCGCGGTGTCACAA 1533
222 lPheSerAspTyr 226
1534 TATTTCGCGCTCTAC 1547
seq_name: /cgn2_6/prodata/2/lna/backfiles1.seq:5244657-4
seq_documentation_block:
; Patent No. 5244657
; APPLICANT: KLIEN, MICHEL A.; BOUX, HEATHER A.; COCKLE,

```

```

;STEPHEN A.; LOOSMORE, SHEENA M.; ZEALY, GAVIN R.
; TITLE OF INVENTION: GENETIC DETOXIFICATION OF PERTUSSIS
; TOXIN
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/589,423
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 275,376
; FILING DATE: 23-NOV-1988
; SEQ ID NO: 4
; LENGTH: 4208
5244657-4

```

```

alignment_scores:
Quality: 89.00 Length: 221
Ratio: 1.000 Gaps: 11
Percent Similarity: 40.271 Percent Identity: 23.529

```

```

alignment_block:
US-09-528-682-1 x 5244657-4 ..

```

```

Align seq 1/1 to: 5244657-4 from: 1 to: 4208

```

```

5 LeuTyrArgAlaAspSerArgProProAspGluIleLysArgSerGly 21
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
627 GTATACCGCTATGACTCCCGCGCGGAGAGAGCTTTCCAGAAC...GG 673
21 yLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMet 38
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
674 ATTCACGCGCTGGGGAACAACAGAC..... 698
38 snlLeuAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
||||| :::: ||| ::::
699 ....AATGCTCGAACATCTGACCGGACGCTTCCTCCAGTCCGCGAGC 743
55 TyrAspAspGlyTyrValSerThrSerLeuSerLeuArg..... 67
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
744 AGCAACAGCGCTTCTGCTCCACGACGACGCGCGGCTATACGAGGT 793
68 .....SerAlaHisLeuAlaGly 74
794 CTATCTGAAACATCGCATGCAGAAAGCGGTCGAGCGGCGGCGGCA 843
74 InsertIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAlaThr 90
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
844 GGGGC.....ACGGCCACTTTCATCGGCTACATTCAGAAATC...CGC 884
91 AlaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerProH 107
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
885 GCCGACAAACAATTTC..... 899
107 sProTyrGlnGlnValSerAlaLeuGlyIleProTyrSerGlnI 124
899 ..... 899
124 lEtyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArgLeuHis 140
||||| |||
900 ..TACGGCGCGCGGAGCTCGTACTTC..... 923
141 ArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn.....I 155
||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
924 .....GAATACGTCGACACTTATGGGACAAATGCGCGGTATCT 964
155 eAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisG 172
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
965 CGCGCGCGCG.....CTGGCCACTTACCAAGCAATATC 999
172 lAlaTyrArgGluGluProTPrIleHisAlaProGlnGlyCysGly 188
||| ||||| |||
1000 TGGCACACGCGGCGCATTCG.....CCGGAACAATCCGC 1034

```



```
189 AsnSerArgThr.....IleThrGlyAspThrCysAsnG1 201
::: |||::: |||::: |||
1035 AGGTAACCGCGGGTCTATGCACACGGCATCACCGCGGAGACCAACGACGAC 1084
201 uGluThrGlnAsn 205
||| ::|||
1085 GGAGTATTCACAC 1097

seq_name: /cgn2_6/ptodata/2/1na/backup1.seq:5433945-4

seq_documentation_block:
; Patent No. 5433945
; APPLICANT: KLEIN, MICHEL H.; BOUX, HEATHER A.; COCKLE,
; STEPHEN A.; LOOSMORE, SHEENA M.; ZEALEY, GAVIN R.
; TITLE OF INVENTION: IMMUNOPROTECTIVE GENETICALLY-DETOXIFIED
; MUTANTS OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/979,798
; FILING DATE: 20-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 589,423
; FILING DATE: 28-SEP-1990
; APPLICATION NUMBER: 275,376
; FILING DATE: 23-NOV-1988
; SEQ ID NO: 4
; LENGTH: 4208
5433945-4

Alignment_scores:
Quality: 89.00 Length: 221
Ratio: 1.000 Gaps: 11
Percent Similarity: 40.271 Percent Identity: 23.529

Alignment_block:
US-09-528-682-1 x 5433945-4 ..

Align seg 1/1 to: 5433945-4 from: 1 to: 4208

5 LeuTyrArgAlaAspSerArgProPAspGluIleLysArgSerGlyG1 21
::: |||::: |||::: |||::: |||::: |||::: |||
627 GATATACCGCTATGACTCCCGCGCGGAGGACGTTTCCAGAC...GG 673
21 yLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMetA 38
::: |||::: |||::: |||
674 ATTCAAGCGCGTGGGAAACACGAC..... 698
38 snIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||::: |||::: |||
699 ....AATGTGCTCGAACATCTGACCGGACGTTCTGCCAGTCGGCAGC 743
55 TyrAspAspGlyTyrValSerThrSerLeuSerLeuArg..... 67
::: |||::: |||::: |||
744 ACCAACAGCGCTTGTGCTCCACACGACGACCGCGCTATACGAGGT 793
68 .....SerAlaHisLeuAlaGlyG 74
794 CTAATCGAACATCCATGACAGAACGCGTCCAGCGCAACCGCGCGCA 843
74 InSerIleLeuSerGlyTyrSerThrTyrIleTyrValIleAlaThr 90
::: |||::: |||::: |||
844 GGGGCG.....ACCGGCACTTCATCGCTACATCTACGAAAGTC...CGC 884
91 AlaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerProH1 107
||| ||| |||
885 GCCGACAAACAATTTC..... 899
107 sProTyrGluGlnGluValSerAlaLeuGlyIleProTyrSerGlnI1 124
899 ..... 899
124 LeTyrGlyTyrPtyrArgValAsnPheGlyValIleAspGluArgLeuHis 140
```

```
900 ..TACGGCCCCCGACCTGTAATTC..... 923
||| |||
141 ArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn.....11 155
||| ||| |||
924 .....GAATACGTGACACTTATGGCAACATGGCGCGTATCTCT 964
155 eAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAsnHisG 172
||| ||| |||
965 CGCGCGCGCG.....CTGGCACTTACACAGAGCGGAATATTC 999
172 InAlaTyrArgGluGluProTyrPheHisAlaProGlnGlyCysGly 188
||| ||| |||
1000 TGGCAACCGCGCGCATTCG.....CCCAAAACATCCCG 1034
189 AsnSerArgThr.....IleThrGlyAspThrCysAsnG1 201
::: |||::: |||::: |||
1035 AGGTAACCGCGGGTCTATGCACACGGCATCACCGCGGAGACCAACGACGAC 1084
201 uGluThrGlnAsn 205
||| ::|||
1085 GGAGTATTCACAC 1097

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-103-840A-1

seq_documentation_block:
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Kv
US-09-103-840A-1

Alignment_scores:
Quality: 94.50 Length: 230
Ratio: 1.086 Gaps: 13
Percent Similarity: 37.826 Percent Identity: 24.348

Alignment_block:
US-09-528-682-1 x US-09-103-840A-1/rev ..

Align seg 1/1 to reverse of: US-09-103-840A-1 from: 1 to: 4411529

6 TyrArgAla.....AspSerArgProPAspGluIleLysArgse 19
||| ||| |||
3735564TATCGAGGCCATCCACATCTCCGACATCCCGATCAACATCAAGAA 3735515
19 r.....GlyGlyL 22
3735514CCCTTACCTCGGACCTTGTCGTCGCCCATCAAGCTCCGCGCAATC 3735465
22 eumetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsn 38
||| ||| |||
3735464TGCCTCGCGGTGCGCATTACCAATATTTGATCGCCCATCA..... 3735421
39 IleAsnLeuTyrAspHis...AlaArgGlyThrGlnThrGlyPheValAr 54
||| ||| |||
3735420....AATCAATCCCATCACCTGTGGCGGACAA..... 3735388
```



```

54 gtyraspspglytyrvalserthrserleuserleuargseralaHisL 71
3735387.....CTTCACCAACCAACATC 3735372
71 euAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrIleTyrVal 87
3735371ACGTGGCTGGCCCGCTCTC.....GTC 3735349
88 IleAlaThrAlaProAsn..... 93
3735348CATTAACAATTCACCAATCCAAAGTCGCCCTCAGCCCTTCGCCCATTC 3735299
94 .....MetPheAsnValA 98
3735298CCACACCTGATCGGCCCAATACATATCAATACCGGGTTTCCATCCCA 3735249
98 snAspVal...LeuGlyValTyrSerProHisProTyrGlnGlnVal 113
3735248GTAACTTCAGTTATTCACCCGACCCCTTTCCTCC..... 3735208
114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTrpTyrArgVa 130
3735207.....GGTTGG..... 3735202
130 IAsnPhgGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArgA 147
3735201.....CCACACATTCACCGGGGGGGCGCTACCCG 3735168
147 sPArGTyTyrArgAsnLeuAsnIleAlaProAlaGlnAspGlyTyrArg 163
3735167ACCTCT...TGCGGTAAAG...CAGGACCGAGGCGCTTCACCAATTCGCG 3735125
164 LeuAlaGlyPheProProAspHisGlnAlaTrpArg_GluGluProTrpI 180
3735124GATTCCTCGATTCGAGGACCAACCATTCGCGTGGCATCAACGTGATTGGC 3735075
180 LeHisHisAlaProGlnGlyCysGlyAsnSerSerArg 192
3735074CACATCAACGCCCTGAGCACCCCGGCATACACATCGA 3735037
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-251-645-1
seq_documentation_block:
: Sequence 1, Application US/09251645
: Patent No. 6281413
: GENERAL INFORMATION:
: APPLICANT: Kramer, Vance C.
: APPLICANT: Morgan, Michael K.
: APPLICANT: Anderson, Arne R.
: APPLICANT: Hart, Hope
: APPLICANT: Warren, Gregory W.
: APPLICANT: Dunn, Martha
: APPLICANT: Chen, Jeng S.
: TITLE OF INVENTION: NOVEL INSECTICIDAL TOXINS FROM PHOTORHABDUS LUMINESCENS
: FILE REFERENCE: CSCI963/A
: CURRENT APPLICATION NUMBER: US/09/251,645
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 9717
: TYPE: DNA
: ORGANISM: Photorhabdus luminescens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (412)..(1665)
: OTHER INFORMATION: orf1 ~46.4 kda
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1666)..(2447)
: OTHER INFORMATION: orf2 ~28.1kda

```

```

: FEATURE:
: NAME/KEY: CDS
: LOCATION: (2758)..(3318)
: OTHER INFORMATION: orf3 ~20.7 kda
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (3342)..(4118)
: OTHER INFORMATION: orf4 ~28.7 kda
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (4515)..(9269)
: OTHER INFORMATION: orf5 ~176 kda
US-09-251-645-1

alignment_scores:
  Quality: 88.00      Length: 114
  Ratio: 1.419      Gaps: 4
  Percent similarity: 54.386      Percent identity: 26.316

alignment_block:
US-09-528-682-1 x US-09-251-645-1 ..
Align seg 1/1 to: US-09-251-645-1 from: 1 to: 9717
27 HisAsnGluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAs 43
8142 ATTCACACAGCATATGCTTCGCGTAAAGTACGACGATATTCATTGGGGA 8191
43 PHisAlaArgGlyThrGlnIleThrGlyPheValArgTyrAspAspGlyTyrV 60
8192 TGAACG..... 8198
60 aIserThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIle 76
8199 ..ATTACTGTGGTTCAGCGATGATTAAGCAACGATTTGACATCAAGGCC 8246
77 LeuSerGlyTyrSerThrTyrTyrIleTyrValIle.....AlaTh 90
8247 AAGATGTGTCATCATGCTACATATCATATGATTTGACACATACAGTAC 8296
90 rAlaProAsnMetPheAsnValAsnAspValLeuGlyVal...TyrSerP 106
8297 GGTATTCGCCAGTCCAGAACGAAACGAGTTGATGGCTTATCTCTTACAC 8346
106 rHisProTyrGlnGlnGlnValSerAlaLeuGlyGlyIleProTyrSer 122
8347 CTTATGGCTTTAGAGTTTAATTTCTCATTTACCGGGTTTGATGGCGCA 8396
123 Gln.....IleTyrGlyTyrTyrArgValAsnPhgGly 133
8397 CAGGTTGATTCAGTAAACAGCGCTGATCTTCAAGTAAACGGA 8438
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-296-848A-1
seq_documentation_block:
: Sequence 1, Application US/08296848A
: Patent No. 6019982
: GENERAL INFORMATION:
: APPLICANT: Clements, John D.
: APPLICANT: Dickinson, Bonny L.
: TITLE OF INVENTION: MUTANT ENTEROTOXIN EFFECTIVE AS A
: TITLE OF INVENTION: NON-TOXIC ORAL ADJUVANT
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PENNIE & EDMONDS LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```



```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,848A
FILING DATE: 26-AUG-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5113-046
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNTE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-296-848A-1

alignment_scores:
Quality: 86.00 Length: 15
Ratio: 5.733 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-296-848A-1 ..

Align seq 1/1 to: US-08-296-848A-1 from: 1 to: 45
186 G1CysG1yAnSerSerArgThrIleThrG1yAspThrCysAsn 200
|||||
1 GGTTGCGAATTCATCAAGACATTCACGTCATCTGTAT 45

seq_name: /cgn2.6/ptdata/2/1na/5A.COMB.seq:US-08-204-656B-9

seq_documentation_block:
Sequence 9, Application US/08204656B
Patent No. 5538882
GENERAL INFORMATION:
APPLICANT: Matsui, Ikuro
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
TITLE OF INVENTION: Oligosaccharide Using The Enzyme
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
```

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
HYPOHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
LIBRARY: Library of chromosomal DNA from Bacillus
LIBRARY: macerans, pMAC, generated by treating chromosomal DNA from Bacillus
LIBRARY: IM1243 with a restriction enzyme, and inserting and linking restric
LIBRARY: fragments to pBR322
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2061
OTHER INFORMATION: /note= "Nucleotides 1-2061
OTHER INFORMATION: correspond to nucleotides 82-2142 of the Bacillus macerans
OTHER INFORMATION: cyclomalodextrin glucanotransferase structural gene"
US-08-204-656B-9

alignment_scores:
Quality: 83.50 Length: 241
Ratio: 0.852 Gaps: 13
Percent Similarity: 40.664 Percent Identity: 21.577

alignment_block:
US-09-528-682-1 x US-08-204-656B-9 ..

Align seq 1/1 to: US-08-204-656B-9 from: 1 to: 2061
18 ArgSerG1yLeuMetProArgG1yHisAsnG1yThrPheAspArgG1 34
|||||
391 AGGTGCGATCGACTTCGCCCCCAACACAC..... 420
34 yThrG1MetAsnLeuAsnLeuTyAspRHisAlaArg1yThrG1nThrG 51
|||||
421 .....GTCTCGCGCCGACGACAGGCGACGACCCCG 448
51 lypheValArg.....TyAspAspG1yTyValSerThrSer 63
|||||
449 GCTTCGCCGACAGACGGTCGCGCTGTATGATACGGT..... 483
64 leuSerLeuArgSerAlaHisLeuAlaG1yClnSerLeuSerG1yTy 80
|||||
484 .....TCGCTGCTCGCGCGCTA 500
80 rSerThyTyrTyrIleTyValIleAlaThrAlaProAsnMetPheAsn 97
|||||
501 CAGC..... 504
97 AlaAsnAspValLeuG1yValTySerProHisProTyG1nG1nVal 113
|||||
505 ..AATGATACGGCGCGCTTTTCATCATAC..... 534
114 SerAlaLeuG1yG1yIleProTySerG1nIleTy.....G1yTyrTyAr 129
|||||
535 .....GGGGGACCGGATTTTCCACAGTGTGACAGCGGTATTATCA 575
129 gValAsnPhcIyValIleAspArgLeuArgLeuHisArgAsnArgG1y 146
|||||
576 GAACCTTAGCAGCTGGCGGAC.....ATCAGCATTAACAACAACGCTA 619
146 rGAspArgTyTyrArg.....AsnLeuAsnIleAlaProAlaG1u 159
|||||
620 TCGACGCTTATTATAAAGCGGTATGACGCTTGGCTGGGATGGGTGTG 669
```



```

160 AspGlyTyrArgLeuAlaGly..... 166
161 ||||| ||||| |||||
670 GACGGATTCGTTTGGACGGGTGAACGATATCTTCGCTGGCAAA 719
167 .....PheProAspHisGlnAlaTrpArgGlu 177
170 AAGCTTCGTTCCGATTTACGGCGGATACCGGATTTACGTTG 769
177 LuProTrpLeuHisAlaProGlnGlyCysGlyAsnSerArgThr 193
178 ||||| ||||| |||||
770 GGAATGGTATCTTGGCGGATCAACCGACGAGACACATTAA... 816
194 IleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
817 .....TTGCCACACGAAGCGGATGAACCTGCTGGACTTT.. 852
210 rLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAsp...T 226
853 .....GAATACGGCGAGAAAGTG...CGCGAAGTGTTCGGGACAAA 892
226 yrgInSerGluValAspIleTyr 233
893 CGGAACGATGAGGATCTCTAT 915
seq_name: /cgn2_6/plodata/2/lna/5A.COMB.seq:us-08-470-702-5

```

```

seq documentation block:
Sequence 5, Application US/08470702
Patent No. 5631149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2061 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)

```

```

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-470-702-5

```

```

alignment_scores:
Quality: 83.50 Length: 241
Ratio: 0.852 Gaps: 13
Percent Similarity: 40.664 Percent Identity: 21.577

```

```

alignment_block:
US-09-528-682-1 x US-08-470-702-5 ..

```

```

Align seg 1/1 to: US-08-470-702-5 from: 1 to: 2061

```

```

18 ArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArgI 34
||| ||| ||| |||
391 AGGTCCGATCGACTTCGCCCAACACAC... 420
34 yThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrG 51
421 .....GTCTCCGGCCGACGACGAGGACGACCCG 448
51 LyrPheValArg.....TyrAspAspGlyTyrValSerThrSer 63
449 GCTTCGCCGACGACGCGTGCCTGTATGATAACGT..... 483
64 LeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyr 80
484 .....TTGGCTGCTGGCGGCTTA 500
80 rSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsnV 97
501 CAGC..... 504
97 AlaAsnAspValLeuGlyValIleTyrSerProHisProTyrGluGlnVal 113
505 ..AATGATACGCGCGCCCTTTCATCATAC..... 534
114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyr...GlyTyrTyr 129
535 .....GGGGGACCGATTTTCCAGATGTAAGACGATATTTCAA 575
129 gValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyr 146
576 GAACCTGACGACCTGGCGGAC.....ATCAGCATTAACAGACGCTA 619
146 rGAspArgTyrTyrArg.....AsnLeuAsnIleAlaProAlaGlu 159
620 TGAACGCTTATTAAAGCGGTATGACGCTTGGCGATGGGTG 669
160 AspGlyTyrArgLeuAlaGly..... 166
670 GACGGATTCGTTTGAACGGGTGAACGATATCTTCGCTGGCAAA 719
167 .....PheProAspHisGlnAlaTrpArgGlu 177
720 AAGCTTCGTTCCGATTTACGGCGGATACCGGATTTACGTTG 769
177 LuProTrpLeuHisAlaProGlnGlyCysGlyAsnSerArgThr 193
770 GGAATGGTATCTTGGCGGATCAACCGACGAGACACATTAA... 816
194 IleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
817 .....TTGCCACACGAAGCGGATGAACCTGCTGGACTTT.. 852
210 rLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAsp...T 226
853 .....GAATACGGCGAGAAAGTG...CGCGAAGTGTTCGGGACAAA 892
226 yrgInSerGluValAspIleTyr 233
.....

```


893 CGGAACGATGAAGATCTCTAT 915

seq_name: /cgn2_6/ptodata/2/lna/5A.COMB.seq:US-08-467-831-5

seq_documentation_block:

; Sequence 5, Application US/08467831
; Patent No. 5635378

; GENERAL INFORMATION:

; APPLICANT: MATSUI, IKUO

; APPLICANT: ISHIKAWA, KAZUHIKO

; APPLICANT: MIYAIRI, SACHIO

; APPLICANT: HONDA, KOICHI

; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH

; STREET: 8110 Gatehouse Road, Suite 500 East

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22042

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/467,831

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/204,656

; FILING DATE: 02-MAR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: WEINER, MARC S.

; REGISTRATION NUMBER: 32,181

; REFERENCE/DOCKET NUMBER: 234-252P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 205-8000

; TELEFAX: (703) 205-8050

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2061 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (synthetic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-467-831-5

alignment_scores:

Quality: 83.50 Length: 241

Ratio: 0.852 Gaps: 13

Percent Similarity: 40.664 Percent Identity: 21.577

alignment_block:

US-09-528-682-1 x US-08-467-831-5 ..

Align seq 1/1 to: US-08-467-831-5 from: 1 to: 2061

```

18 ArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrRheAspArgG1 34
||||| ||| ||| ||| |||
391 AGTCGGATGACTTCCGCCCAACACAC..... 420
34 yThG1MetAsn1LeuAsnLeuTyrAspHisAlaArgGlyThrGlnThrg 51
::: ||| ||| ||| |||
421 .....GTCTCCGGCCGACGAGGACGAAACCCG 448

```

```

51 LyrPheValArg.....TyrAspAspGlyTyrValSerThrSer 63
||||| ||| ||| ||| |||
449 GCTTCGCCGAGAACGTCGCGTATATACGAT..... 483
64 LeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTy 80
||||| ||| ||| ||| |||
484 .....TCGCTGCTCGCGCGCTA 500
80 rSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsn 97
|||||
501 CACG..... 504
97 AlaAsnAspValLeuGlyValTyrSerProHisProTyrGluGlnGluVal 113
||||| ||| ||| ||| |||
505 ..AATGATACGCGCGCGCTTTCATCATAC..... 534
114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyr...GlyTyrTyr 129
||||| ||| ||| ||| |||
535 .....GGGGGACCGAATTTTCCACGATGACAGCGGATTTTCAA 575
129 gValAsnRheGlyValIleAspGluAlaGlyLeuHisArgAsnArgGlyTyr 146
::: ||| ||| ||| ||| |||
576 GAACCTCTACGACCTGCGGAC.....ATCACCATACACACACCTA 619
146 rGAspArgTyrTyrArg.....AsnLeuAsnIleAlaProAlaGlu 159
||||| ||| ||| ||| |||
620 TGGACGCTTATTTTAAAGCGGATGACCTTGGCTGGAGTGGGTG 669
160 AspGlyTyrArgLeuAlaGly..... 166
||||| ||| ||| ||| |||
670 GACGGAGTGTCTTTTACGCGGCGTGAAGAGTATCTTGGCTGCAAA 719
167 .....PheProProAspHisGlnAlaThrArgGlu 177
||||| ||| ||| ||| |||
720 AAGCTGCTTCTCGATTTACGCGGCGGATATCTTACGTTG 769
177 LyrProTrpIleHisAlaProGlnGlyGlyAsnSerSerArgThr 193
||||| ||| ||| ||| |||
770 GGGATGGTATCTTGGCGGCGGATCAACCGAGACACACATTTAA... 816
817 .....TTCGCCAACGAAGCGGATGAACCTGCGGACTTT... 852
194 IleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTy 210
||||| ||| ||| ||| |||
210 rLeuArgGlyTyrGlnSerLysValLysArgGlnIlePheSerasp...T 226
||||| ||| ||| ||| |||
853 .....GAATACCGCGACGAGATG...CGGAGATGTTCCGGGACAAA 892
226 yGlnSerGluValAspIleTyr 233
||||| ||| ||| ||| |||
893 CGGAACGATGAAGATCTCTAT 915

```

seq_name: /cgn2_6/ptodata/2/lna/5A.COMB.seq:US-08-195-152-1

seq_documentation_block:

; Sequence 1, Application US/08195152

; Patent No. 5679541

; GENERAL INFORMATION:

; APPLICANT: Bonini, Nancy M.

; APPLICANT: Leiserson, William M.

; TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hobach, Test, Albritton & Herbert

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: United States

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk


```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,152
; FILING DATE: 14-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tregeath, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59551/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; US-08-195-152-1

```

```

alignment_scores:
  Quality: 82.50      length: 269
  Ratio: 0.833       Gaps: 14
  Percent Similarity: 36.803   Percent Identity: 20.818

```

```

alignment_block:
US-09-528-682-1 x US-08-195-152-1 ..

```

```

Align seg 1/1 to: US-08-195-152-1 from: 1 to: 3231

```

```

7  ATGAlAspSerArgProProaspGluIleLysArgSerGlyLysLeu 23
   : : : : : : : : : : : : : : : : : : : : : : : :
1002 AAGTCGGAGTCCCGCAGCCCGGCGCAAGTCGACGATCG..... 1040
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23  tProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsnIle 40
   : : : : : : : : : : : : : : : : : : : : : : : :
1041 .....CTGGACAGAGGCTCGGTGGCGGATCCA 1068
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
40  subLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArgTyrAsp 56
   : : : : : : : : : : : : : : : : : : : : : : : :
1069 ATTTGATC.....GGCTGCAGCTCGGCGCAGCATCGCGTGAC 1106
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57  AspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisLeuIle 73
   : : : : : : : : : : : : : : : : : : : : : : : :
1107 GGAGGAGCAGAGTGGCGGTCAACTCTTCGGCAGTGGCAGCG..... 1150
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
73  yGlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAla 90
   : : : : : : : : : : : : : : : : : : : : : : : :
1151 AGCAGCGGTCTACGACGCGCAACATGACTACTACTACTAC..... 1190
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
90  hValIleProAsnMetPheAsnValAlaAspValLeuGlyValTyrSerPro 106
   : : : : : : : : : : : : : : : : : : : : : : : :
1191 .....AACAGCATGCAGCAGCATGACGCGCG 1214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107  HisProTyrGluGlnGlnValSerAlaLeuGlyIleProTyr..... 121
   : : : : : : : : : : : : : : : : : : : : : : : :
1215 CCGCCCTTCTACTCCGATAC.....GGAACCTCTTATCGGCG 1252
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 ..... 121
1253 GCGAAGCGGCGCGCAGCGCAAGATGCAACCGGAGCGCGAGCTGCGG 1302
122 .....Ser 122
1303 CGGCTGCTACTTGAAGCCAGCTATGCGCGCGGCAACACAACTCG 1352
123 GlnIleTyr..... 125
1353 CAGCTGTACAGCAGTCCGTACGCGGCTACACAACTTCGCGGACAGGA 1402

```

```

126 .....GlyTyrTyrArgValAsnPheGly.ValIleAspGluArg... 138
   : : : : : : : : : : : : : : : : : : : : : : : :
1403 CTACGGCGGCTACTACACAGCAGTACGCGACACTTTCACGCGGCCA 1452
139 ...LeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeu 154
   : : : : : : : : : : : : : : : : : : : : : : : :
1453 ACTACTACCCGATGCGGTGCGAGCTCCGCCAGCTCGAGTGGAGTGA 1502
154 nIleAla.....ProIleGlnAspGlyTyrArgGluAla.Gly 166
   : : : : : : : : : : : : : : : : : : : : : : : :
1503 CATGCTTCATGTGCGCGGCTCTCGAATCTCCGAGAGTCCACGGGA 1552
167 PhePro.....ProAspHisGlnAlaTyrArg..... 175
   : : : : : : : : : : : : : : : : : : : : : : : :
1553 CACCACATCGACAGCCGCGGTGCACACAGACACCACACTCGCCGCA 1602
176 .....GluGluProTyrIleHisAlaProGlnGly.... 186
   : : : : : : : : : : : : : : : : : : : : : : : :
1603 CGCTCCGATCTCGCGGACGACTGCTCCGCGCATGCGCGGCGCAT 1652
187 .....CysGlyAsnSerSerArgThrIleThrGlyAspThr 198
   : : : : : : : : : : : : : : : : : : : : : : : :
1653 GTGTCCGCGGACGCTGCGCGCTGCTCTCACTCGAGCGAGGACAGAG 1702
199 Cys 199
   |||
1703 TGT 1705

```

```

seq name: /cgn2_6/ptodata/2/1na/PC/TUS.COMB.seq:PCT-US96-05611A-28

```

```

seq documentation_block:

```

```

; Sequence 28, Application PC/TUS9605611A
; GENERAL INFORMATION:

```

```

; APPLICANT: Mueller, John P.
; APPLICANT: Lenardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott
; APPLICANT: Nye, Steven H.
; APPLICANT: Pelitrey, Clara M.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: Wilkins, James A.
; TITLE OF INVENTION: Modified Myelin Protein Molecules
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
; COMPUTER: Macintosh Centris 610
; OPERATING SYSTEM: System 7
; SOFTWARE: Microsoft Word 6.0.1

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05611A
; FILING DATE: 02-MAY-1995
; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,644
; FILING DATE: May 2, 1995

```

```

; APPLICATION NUMBER: 08/431,648
; FILING DATE: May 2, 1995

```

```

; APPLICATION NUMBER: 08/482,114
; FILING DATE: June 7, 1995

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399

```

```

; REFERENCE/DOCKET NUMBER: ALX-129
; TELECOMMUNICATION INFORMATION:

```



```

147 ASPATG..... 148
148 |||
149 GCGCCGACAGAGCTGCTGAAGATGCTATTGGTGGAGGAAAGTGACTCT 842
149 .TyrTYrArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuA 165
149 |||||:::||||: ||| ||:::
843 CAGGATCCCGGAATGTATGGTTCACAGT...GAAGAGAGTTTTCACCTGC. 888
165 IagLYpheProPheSPHsGln..... 172
165 ||| |||||:::
889 ....TTCTTCGAGATCATCTTACCAAGAGCAGCAGCAATGGAATTG 933
173 .....AlaTyrArgGluGluProThrIleHisAlaI 183
173 |||||:::||||: |||||:::
934 AAGTAGAAGATCCCTTACTGGCTCGAGAGATCCGGGA...CATGAAGC 980
183 aProGInGLyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysA 200
183 | ||| :::: ||| ::::
981 CCTCACTGGCAGCAAGAAAGCTATTGTAGACCTATTCTCCAAAACATACC 1030
200 sngLgLUthrgLnaAsnLeuSerThrIleTYrLeuArgGlu..... 213
1031 AAGACTATGAGTATCTCATCAATGATGATCATGCCTTCACATATGCTGAG 1080
214 .....TyGlnSerIysValIysArgGlnIlePheSerAspTYrGlnSe 228
214 ||| :::: ||||| |||||:::
1081 GCGCTTCAACCCACCGCGCAGTCAGCAATCTTTGGCGACTACAAGAC 1130
228 rGLuVal 230
1131 : ::
1131 CACATC 1137

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-449-045C-23
seq_documentation_block:
Sequence 23 Application US/08449045C
Patent No. 5770203
GENERAL INFORMATION:
APPLICANT: Burnette, Neal W.
APPLICANT: Kaslow, Harvey R.
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
TITLE OF INVENTION: SUBUNIT ANALOGS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,045C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/694,733
FILING DATE: 02-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/271,222
FILING DATE: 06-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196C
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid

```


STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-449-045C-23

alignment_scores:
Quality: 80.00 Length: 15
Ratio: 5.714 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 80.000

alignment_block:
US-09-528-682-1 x US-08-449-045C-23 ..

Align seg 1/1 to: US-08-449-045C-23 from: 1 to: 50

179 TrrlEhIshIsAlProGInGlyCysGIyAsnSerSeraTgr 193
|||||
3 TGGATTCAATCATGCACCGCGGTTGTGGAAATGCTCCAAGATCA 47

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-449-045C-24

seq_documentation_block:

Sequence 24, Application US/08449045C
Patent No. 5770203
GENERAL INFORMATION:
APPLICANT: Burnette, Neal W.
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449, 045C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/694,733
FILING DATE: 02-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/271,222
FILING DATE: 06-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196C
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-449-045C-24

alignment_scores:
Quality: 80.00 Length: 15
Ratio: 5.714 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 80.000

alignment_block:

US-09-528-682-1 x US-08-449-045C-24 ..
Align seg 1/1 to: US-08-449-045C-24 from: 1 to: 50

179 TrrlEhIshIsAlProGInGlyCysGIyAsnSerSeraTgr 193
|||||
3 TGGATTCAATCATGCACCGCGGTTGTGGAAATGCTCCAAGATCA 47

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-35

seq_documentation_block:

Sequence 35, Application US/08435605A
Patent No. 5874287
GENERAL INFORMATION:
APPLICANT: Burnette, W. Neal
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Amgen Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435, 605A
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mazza, Richard J.
REGISTRATION NUMBER: 27,657
REFERENCE/DOCKET NUMBER: A-196B
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "oligonucleotide"
US-08-435-605A-35

alignment_scores:
Quality: 80.00 Length: 15
Ratio: 5.714 Gaps: 0
Percent Similarity: 93.333 Percent Identity: 80.000

alignment_block:
US-09-528-682-1 x US-08-435-605A-35 ..

Align seg 1/1 to: US-08-435-605A-35 from: 1 to: 50

179 TrrlEhIshIsAlProGInGlyCysGIyAsnSerSeraTgr 193
|||||
3 TGGATTCAATCATGCACCGCGGTTGTGGAAATGCTCCAAGATCA 47

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-435-605A-36

seq_documentation_block:

Sequence 36, Application US/08435605A
Patent No. 5874287
GENERAL INFORMATION:
APPLICANT: Burnette, W. Neal
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN


```

: TITLE OF INVENTION: SUBUNIT ANALOGS
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Amgen Inc.
: STREET: 1840 De Havilland Drive
: CITY: Thousand Oaks
: STATE: California
: COUNTRY: USA
: ZIP: 91320-1789
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: US/08/435,605A
: APPLICATION NUMBER: US/08/435,605A
: FILING DATE: 05-MAY-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Mazza, Richard J.
: REGISTRATION NUMBER: 27,657
: REFERENCE/DOCKET NUMBER: A-196B
: INFORMATION FOR SEQ ID NO: 36:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 50 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: /desc = "oligonucleotide"
US-08-435-605A-36

alignment_scores:
  Quality: 80.00 Length: 15
  Ratio: 5.714 Gaps: 0
  Percent Similarity: 93.333 Percent Identity: 80.000

alignment_block:
US-09-528-682-1 x US-08-435-605A-36 ..
Align seg 1/1 to: US-08-435-605A-36 from: 1 to: 50

179 TTTTCTCATCATGCGACGCGAGGGTTGTGGAGATGCTCCAGAGATCA 47
|||||
seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-128-450-25

seq_documentation_block:
: Sequence 25, Application US/09128450
: Patent No. 6211149
: GENERAL INFORMATION:
: APPLICANT: Chesebro, Bruce W
: APPLICANT: Chabrey, Byron W
: APPLICANT: Chabrey, Joelle
: APPLICANT: Priola, Susette
: TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
: FILE REFERENCE: 50121
: CURRENT APPLICATION NUMBER: US/09/128,450
: PRIOR FILING DATE: 1998-08-03
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 25
: LENGTH: 1000
: TYPE: DNA
: ORGANISM: Hamster sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (60)..(824)
US-09-128-450-25

```

```

alignment_scores:
  Quality: 80.00 Length: 124
  Ratio: 1.176 Gaps: 6
  Percent Similarity: 54.839 Percent Identity: 24.194

alignment_block:
US-09-528-682-1 x US-09-128-450-25 ..
Align seg 1/1 to: US-09-128-450-25 from: 1 to: 1000

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrArgVa 130
:|||||
417 GCCCTGGTGGGGGCGCTT.....GGTGGCTACAGTCT 448
:|||||
130 LAsnPhcGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArgA 147
:|||||
449 GGGGAGTGGCCATGACGAGCCCATGATGATTTGGCAATGACTGGGAGG 498
:|||||
147 sPArgTyrTyrArg...AsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
:|||||
499 ACCGCTACTACCGTGAACATGAAACCGCTACCTTAACCAAGTATATAC 548
:|||||
163 ArgLeuAlaGlyPheProAspHisGlnAlaTrrPArgGluGluProTr 179
:|||||
549 CGG.....CCAGTGGACCATGACAAACCAACGAGAACTT 583
:|||||
179 rIleHisHisAlaProGlnGlyCysGlyAsn.....SerSerArgT 193
:|||||
584 TGTCAC.....GATTGTGTACATCATCACCATCAACGACGACA 621
:|||||
193 hrlleThnGlyAspThrCysAsnGlnGluIuThrGlnAsnLeuSerThrlle 209
:|||||
622 CAGTCACCAACCAACCAAGGGGAG.....AACTTACGAGAGACC 662
:|||||
210 TyrLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAspTyr 226
:|||||
663 GACATCAAGATATGAGACCGCGTGTGGAGCAAGATGTGTACACCACTA 712
:|||||
226 rGlnSerGluValAspIleTyr 233
:|||||
713 TCAGAGGAGTCCAGCGCTTAC 734

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-823-494-25

seq_documentation_block:
: Sequence 25, Application US/09823494
: Patent No. 6355610
: GENERAL INFORMATION:
: APPLICANT: Chesebro, Bruce W
: APPLICANT: Chabrey, Byron W
: APPLICANT: Chabrey, Joelle
: APPLICANT: Priola, Susette
: TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
: FILE REFERENCE: 50121
: CURRENT APPLICATION NUMBER: US/09/823,494
: PRIOR FILING DATE: 2001-03-30
: CURRENT APPLICATION NUMBER: 09/128,450
: PRIOR FILING DATE: 1998-08-03
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 25
: LENGTH: 1000
: TYPE: DNA
: ORGANISM: Hamster sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (60)..(824)
US-09-823-494-25

alignment_scores:

```


Quality:	80.00	Length:	124
Ratio:	1.176	Gaps:	6
Percent Similarity:	54.839	Percent Identity:	24.194

Align seg 1/1 to: US-09-823-494-25 from: 1 to: 1000

```

114 SerAlaLeuGlyGlyLeuProTyrSerGlnLeuTyrGlyTryptyrArgva 130
      :::::::::::::::  |||  |||  ::
417 GCCGTGTTGGGGGCCCTT.....GCTGCTACATGCT 448

```

```

130 lasnphegl yValIleaspluarbleuHtsargasnargl uTyrarg 147
    :::: ::::| ::::| ::::| ::::| ::::|
449 GGGAGTGCATGAGCAGGCCATGATGCATTTGGCAATGACTGGAGG 498

```

```

147 spargtYrTyrArg...AsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
|||||
|||||
499 ACCGCTACTACCGTGAACAACATGAAACCGCTACCCCTAACCAAGTGTATTAC 548
|||||
|||||

```

163 ArgLeuAlaGlyPheProAspHisGlnAlaTTPargGluGluProTr 179
 ||| ||| ::::: ::
 549 CGG.....CCAGTGGACCACTACACACCAACGACCAACTT 583

```

179 pLEHISHISAlaProGInGlyCysGlyAsn.....SerSerArgT 193
    ::::| | | | | | | | | | | | | | | | | | | | |
584 TGtGCAC.....GATTGTGTCAACATCACCATCAAGCAGCACA 621

```

```

193 hrlethrnrglyaspthrcysasnglglutthrglnasnleuserthrile 209
    ||::|||      |||      ::::|||      |||:::|:
622 CAGTCACCACCACCACCACCAAGGGGAG.....AACTTCACGGAGACC 662

```

210 TyrLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerAspTy 226
 ::::: :: ||::: ||
 663 GACATCAAGATAATGGAGCGCGTGTTGGAGCAGATGTGTACCACCAGTA 712

226 rGlnserGluValaspIeTyr 233
 ||||::||| ::: |||
 713 TCAGAAGGAGTCCACAGGCCTAC 734

name: /cgn2_6/ptodata/2/lna/6A_COMB.seq;US-08-728-603-14
documentation_block:

GENERAL INFORMATION:
APPLICANT: Cesarman, Ethel

TITLE OF INVENTION: PROTEINS OF KAPOSI'S SARCOMA ASSOCIATED
 TITLE OF INVENTION: HERPESVIRUS
 NUMBER OF SEQUENCES: 19

CONTRACT OR DEDICATED ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester

STATE: NEW YORK
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM.

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Datamath Release #1.0 Version #1.20

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,603
FILING DATE: 10-OCT-1996
CLASSIFICATION: 43E

ATTORNEY/AGENT INFORMATION:
NAME: BRAMAN, SUSAN J.
REGISTRATION NUMBER: 34,103
DEEDREVENUE/DOCKET# NUMBER: 10000/4000

TELECOMMUNICATION INFORMATION:

? TELEPHONE: 716-263-1636
 ? TELEFAX: 716-263-1600
 ? INFORMATION FOR SEQ ID NO: 14
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 4332 base pairs
 ? TYPE: nucleic acid
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: cDNA
 ? US-08-728-603-14

alignment_scores:	
quality:	80.00
Ratio:	0.630
Percent Similarity:	48.289
	Percent Identity: 23.194

US-09-528-682-1 x US-08-728-603-14/rev

Align seg 1/1 to reverse of: US-08-728-603-14 from: 1 to: 4332

```

14 Aspgluilelysargserglycyleuemetproargglyhisasnsluty 30
    ||| :::::||| ||| ||||| ||| :::::
934 GATTGACTGGAAACGCCATGGC.....CCACGGGGACATGGACAG.. 3896

```

30 rRheaspragglyThrGlnMetAsnIleasnLeuTyAspHisAlaArgg 47
 ||| :: |||:
 895GGGTGTTAGGCTGTGGTGCACATGCTGCACATGCCACGGGA 3856

47 L yThrgInthrGlyPheValArGtyrAspSpGlyTyrValSerThrser 63
::: ::::| | | | | |
855 TGGACATCAGAGATGGGTC...TATGATCAGGGCAGCGTGCCCGTC 3809

```

64 LeuSerLeuArgSerAlaHis..... 70
   :: ::|||::|||
808 ACTGATGTAGTCCGCCACCGTGAGTTGCTGTGGGTTCTGGGAT 3759

```

71LeuAlaGlyGln 75
758 AGTGTGGCTGGCAGGGCTCATCCGGCATTTCCATGGTAGTGAGG 3709

75 erlleleuSerGlyTyrSerThrTyrTyrIle....TyrValIleala 89
:: |||::: ||| ::::: :::::
708 GTTATCTCGCCTCGCTGTCTCAGTATGTACTCGAGGCGTCCTGTCGTA 3659

90 ThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerPr 106
 |||:::|||||::: ||| ::|
 558 CCGGACCCCCAGGTACTCTCCCTGGGCCCCAGCT..GGGACGACCCGTC 36122

106 o.....HisProTyrGluGlnGluValSerAlaLeu.GlyGlyIlePro 120
| | | : : : : :
561 CCGGCACACTCGGAGGAAACGGCTTTAGTGTTGTGAGGGATCTGATG 3562

```

121 TyrSerGlnIleTyrGlyTrpTyrArgValAsnPhcGlyValIleAspG1 137
:::|||||  |||  ::::  ::::  ::
TTTAGCCAG.....TGGCTGTCATACAGCTTGGACACGTTGGTCTC 3521

```

```

137 uArgLeuHisArgAsnArg.....GluTyrArgAspArgT 149
      :::::      :::::      |||
220 CAGGTTTACCGCCGACGCGTGGGGTGTGGGTCGCTAC.....G 3480

```

149 yrrTrrArg.AsnLeuAsnIleAlaProAlaGluAspGlyTrrArgLeuAl 165
||| ::::: ::||| :: |||
79 TGATGTGAGGATTCGACCGGCCACTACACCCAGGGCCACCAGCAGC 3430

165 aglyPheProPoaSPHisGlnAlaTrpArgLys..... 176
166 ||||| ||||| ||||| ||||| |||||
167 TGGGAAGCCCACTCGGCACACAGATGGAGAATTGTTCGGTCTGTAG 3380

1177GluProTrpIleHisHis.....AlaPro.Gl 185
 ::: ||| |||::: ||||| ::
3729 AAACCTCTGTCAGGCTGACAGGCACAGCTAGCGGTCGTTACACAGAGCCAGGA 3330


```
185 n.....GlyCysGlyAsnSerSerArgThrIleThrGlyAspThrC 199
      :      |||:||||| :      :      :      :      :      :
3329 CCCATCCCCCTGGCGCTGGCTAGCTGGCTGGCGCTGTTCTAGACAT 3280
      :      :      :      :      :      :      :      :
199 ysaen.....GluGluThrGlnAsnLeuSerThrIleTyrLeu... 211
      |||||      |||      |||:|||||:|||||:|||||      |||
3279 GTAATCAGACCAGAGAACCCCGACAGAGACTGCTCGTTAACTCTTC 3230
      :      :      :      :      :      :      :      :
212 ArgGluTyrGlnSerLysValLysArgGlnIle 222
      :      :      :      :      :      :      :      :
3229 CACAGTCACCGTGGCCACCTCAAGCCCGTGT 3197

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-770-379-20

seq_documentation_block:
; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ. ID NO. 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-770-379-20

alignment_scores:
      Quality: 80.00      Length: 263
      Ratio: 0.630      Gaps: 18
      Percent Similarity: 48.289      Percent Identity: 23.194

alignment_block:
US-09-528-682-1 x US-08-770-379-20 ..
Align seg 1/1 to: US-08-770-379-20 from: 1 to: 32207

14 AspGluIleLysArgSerGlyGlyLeuMetProArgGlyHisangLutry 30
||| :|||:|||||
25289 GATTGACTGGAAAGCCATGGC.....CCAGGGGAGACATGACAG.. 25327
```

```
30 rPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArg 47
      :      |||      :      :      :      :      :
25328 .....GGGTGTAGGTCTGGTGGAGACATGCTGCCACTGCCACGGA 25367
      :      :      :      :      :      :      :      :
47 TyrThrGlnThrGlyPheValArgTyrAspArgGlyTyrValSerThrSer 63
      :      :      :      :      :      :      :      :
25368 TGGAAACATCAGACAGATGGGTC...TATGATCAGGGCAGAGGTCTCGCCGTC 25414
      :      :      :      :      :      :      :      :
64 LeuSerLeuArgSerAlaHis.....LeuAlaGlyGlns 70
      :      :      :      :      :      :      :      :
25415 ACTGGATGTATAGTCCGGCCACCGTGGAGTTCTGTGGGATTTCGGGAT 25464
      :      :      :      :      :      :      :      :
71 .....LeuAlaGlyGlns 75
      :      :      :      :      :      :      :      :
25465 AGTGTCTGGCTGGCAGGCTCATCCCGGCATTTCATGGTAGGTGAGG 25514
      :      :      :      :      :      :      :      :
75 erileLeuSerGlyTyrSerThrTyrTyrIle.....TyrValIleAla 89
      :      :      :      :      :      :      :      :
25515 GTTATCTCGCTCGCTGCTCAGTATGTATGTCGAGGGCGCTCGCTCGTA 25564
      :      :      :      :      :      :      :      :
90 ThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerP 106
      :      :      :      :      :      :      :      :
25565 CCGGACCCCGCAGGACTCTCCCTGGGCCACACT...GGGCAGCACCGTCC 25611
      :      :      :      :      :      :      :      :
106 o.....HisProTyrGluGlnGluValSerAlaLeu GlyGlyIlePro 120
      :      :      :      :      :      :      :      :
25612 CCGGCACACTCGGAGGAAAGCGCTTGTAGTGTCTGAGGAGTCTGATG 25661
      :      :      :      :      :      :      :      :
121 TyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGlyValIleAsp 137
      :      :      :      :      :      :      :      :
25662 TTTAGCCAG.....TGGCTGTATACACAGCTTGGACAGTGTGTCTC 25702
      :      :      :      :      :      :      :      :
137 uArgLeuHisArgAsnArg.....GluTyrArgAspArgT 149
      :      :      :      :      :      :      :      :
25703 CAGGTTTACCCGCCAGCGCTGGGTGTGTGTGTCCTGAC.....G 25743
      :      :      :      :      :      :      :      :
149 yTyrTyrArg.AsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAl 165
      :      :      :      :      :      :      :      :
25744 TGTATGTGAGGAGTCCGACCGGCCACTACACCCAGGCCACACGACGAC 25793
      :      :      :      :      :      :      :      :
165 aGlyPheProProAsnHisGlnAlaTPrArgGlu..... 176
      :      :      :      :      :      :      :      :
25794 TGGAAAGCCACCTCGCCACACAGATGGAGATGTGTGGTCTGTTAG 25843
      :      :      :      :      :      :      :      :
177 .....GluProTyrIleHis.....AlaPro.Gl 185
      :      :      :      :      :      :      :      :
25844 AAATCTGTACAGGTGGAGGACACAGGTAGGCTGTTACACAGCGCCAGGA 25893
      :      :      :      :      :      :      :      :
185 n.....GlyCysGlyAsnSerSerArgThrIleThrGlyAspThrC 199
      :      :      :      :      :      :      :      :
25894 CCCATCCCCCTGGCGCTGGCTAGCTGGCTGGCGCTGTTCTAGACAT 25943
      :      :      :      :      :      :      :      :
199 ysaen.....GluGluThrGlnAsnLeuSerThrIleTyrLeu... 211
      |||||      |||      |||:|||||:|||||:|||||      |||
25944 GTAATCAGACCAGAGAACCCCGACAGAGACTGCTCGTTAACTCTTC 25993
      :      :      :      :      :      :      :      :
212 ArgGluTyrGlnSerLysValLysArgGlnIle 222
      :      :      :      :      :      :      :      :
25994 CACAGTCACCGTGGCCACCTCAAGCCCGTGT 26026

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-757-669A-20

seq_documentation_block:
; Sequence 20, Application US/08757669A
; Patent No. 6183751
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
```


Align seg 1/1 to: US-09-230-371A-20 from: 1 to: 32207


```

; FILING DATE: 19920519
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB90/02031
; FILING DATE: 28-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8929293.2
; FILING DATE: 29-DEC-1989
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1948 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Clostridium difficile
; US-07-849-438-1

```

```

alignment_scores:
      Quality: 78.50      Length: 159
      Ratio: 1.106      Gaps: 10
      Percent Similarity: 44.654      Percent Identity: 24.528

```

alignment_block:

US-09-528-682-1 x US-07-849-438-1 ..

Align seg 1/1 to: US-07-849-438-1 from: 1 to: 1948

```

26 GlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsnLeuTy 42
   |||  ::  ||| |||  |||  |||  |||  |||  |||  |||
979 GGTAAATAATATTAATCTT.....AATACACACTGC 1010
   ::  |||  ||| ||| ||| ||| ||| ||| ||| |||
42 rAspHisAlaArgGlyThrGlnThrGlyPheValAlaGlyTrspsGly 58
   ::  |||  ||| ||| ||| ||| ||| ||| ||| |||
1011 TGAAGCAGCTGATGATGCGAACT.....ATTGATGGTA 1045
   ::  |||  ||| ||| ||| ||| ||| ||| ||| |||
59 .....TyrValSerThrSerLeuSerLeuArgSerAlaHisLeuAla 72
   |||  |||  ||| ||| ||| ||| ||| ||| |||
1046 AAAAATATTAATTAATCTAATACACTGCTATAGCTCA.....ACT 1086
   |||  |||  ||| ||| ||| ||| ||| ||| |||
73 GlnGlnSerIleLeuSerGlyTyrSerThrTyrTyr.....I 85
   |||  |||  ||| ||| ||| ||| ||| ||| |||
1087 GGTATATCAATTAATTAATGTAACAATTTTATTTATACATGATGAT 1136
   |||  |||  ||| ||| ||| ||| ||| ||| |||
85 eTyValIleAlaThrAlaProAsnMetPhe..... 95
   |||  |||  ||| ||| ||| ||| ||| ||| |||
1137 TATCGAGAGTGTAAAGGACCTAATGATTTGAATATTTCACCTGCTA 1186
   |||  |||  ||| ||| ||| ||| ||| ||| |||
96 .....AsnValAsnAspValLeuGlyValTyrSerProHisProTyrGlu 110
   |||  |||  ||| ||| ||| ||| ||| ||| |||
1187 ATACTGATGCTAATATACATAGAGGT.....CAAGCTATAGTTATCA 1230
   |||  |||  ||| ||| ||| ||| ||| ||| |||
111 GlnGlnValSerAlaLeuGlyGlyIleProTyr..... 121
   |||  |||  ||| ||| ||| ||| ||| ||| |||
1231 AGTAATTTTAACTTGATGATGTAATAAATATTAATCTTGTATTAATCTC 1280
   |||  |||  ||| ||| ||| ||| ||| ||| |||
122 SerGlnIleTyrGlyTyrPyrGlyValAsnPheGlyValIleAspGlu 138
   |||  |||  ||| ||| ||| ||| ||| ||| |||
1281 AAAAGCAGTACTGATGCGCAT.....ACTATGATAGTACATA 1315
   |||  |||  ||| ||| ||| ||| ||| ||| |||
138 rGleuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
   |||  |||  ||| ||| ||| ||| ||| ||| |||
1316 AA.....AAATATTACTTTAATATACATAC 1338
   |||  |||  ||| ||| ||| ||| ||| ||| |||
155 IleAlaProAlaGluAspGlyTyrArg 163
   |||  |||  ||| ||| ||| ||| ||| ||| |||
1339 ACTGCTGAACGAGCTACTGATGCGCAA 1365

```

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-077-939-16

seq_documentation_block:

```

; Sequence 16, Application US/08077939
; Patent No. 5521088
; GENERAL INFORMATION:
; APPLICANT: FUJII, Toshio
; APPLICANT: IYAMATSU, Akihiro
; APPLICANT: YOSHIMOTO, Hiroyuki
; APPLICANT: MINETOKI, Toshitaka
; APPLICANT: BOGAKI, Takayuki
; APPLICANT: NAGASAWA, Naoshi
; TITLE OF INVENTION: ALCOHOL ACETYLTRANSFERASE GENES AND USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/077,939
; FILING DATE: 18-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 184328/1992
; FILING DATE: 18-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 62997/1993
; FILING DATE: 26-FEB-1993
; ATORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/101 KYPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 346..1923
; US-08-077-939-16

```

alignment_scores:

```

      Quality: 78.00      Length: 208
      Ratio: 0.743      Gaps: 10
      Percent Similarity: 50.481      Percent Identity: 22.115

```

alignment_block:

US-09-528-682-1 x US-08-077-939-16 ..

Align seg 1/1 to: US-08-077-939-16 from: 1 to: 1974

```

68 SerAlaHisLeuAlaGlnSerIleLeuSerGlyTyrSerThrTyrTy 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
528 TCAGCTCACATTAAGCTTGAAGGAATCTGCGAAAAATCCAGCTTT 577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 rIleTyr.ValIleAlaThrAla...ProAsnMetPheAsnValAsnAsp 99
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 TACATATTGTTCTACCAACAAGATGCGCAATATCATGAATTTTATCGC 627
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 ValLeuGlyValTyrSer...ProHisProTyrGlnGlnGlnValSerAl 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
628 AGTCCGAATACATATTCACGGCCACATCCAGTCATGATATATCTCAGT 677
115 a.....LeuGlyIleProTySerGlnIleTyGlyT 127
:|||||:
678 ATTCAAGAATGAAATGAGTGTGCTTCATGAAACAACCTGAGT 727
:|||||:
127 rPYrArGValAsnPhenGlyValIleAspGluArgLeuHisArgAsnArg 143
:|||||:
728 ACAGTCGAGTAATGAAGCAATATTGAAAGAAATTCAAAATAGTAGCGGT 777
144 GluTyArGAspArgTyTyArG.....AsnLeuAsnIle..... 155
:|||||:
778 TCCTATACAGCAAAATTTTAACTTACTACACACTTGCATTCCTTA 827
156 ....AlaProAlaGluAspGlyTyTyArGLeuAlaGlyPheProProAspH 171
:|||||:
828 CTTTGGACCAACAGACCGAGTGGCGGCTAATTGCTCTTCAGAGAGAC 877
171 ts...GlnAlaTrpArgGluGluProTrpIleHisHisAlaProGlnGly 186
:|||||:
878 ACACAGAAAAGTGAGCAAAATTTATCTTGTATCTATATCATTCAGATGCT 927
187 CysGlyAsnSerSerArgTrpTrpIleThrGlyAspTrpCysAsnGluGluH 203
:|||||:
928 GATGGTCGCTCTTCGATCCACTTTTTCATGAT...TTAAGAGAGCAAT 974
203 rGlnAsnLeuSerThr.....IleTyTyLeuArg 213
:|||||:
975 AATATATATTAATAACACACACAAAATAATAGATTCATTTTCACACTAG 1024
213 lUTyGlnSerValLysValArgGlnIlePheSerAspTyGlnSerGlu 229
:|||||:
1025 AGGAGATTCACCAATTATGAGAACTTCAGAACCGATCGAAAAGGTG 1074
230 ValAsp..... 231
:|||||:
1075 ATAGACTTATGACACACCGTACTGTTTATTCGACAGTCACTTCTCGGG 1124
232 ....IleTyAsnArgIleArg 237
:|||||:
1125 TTTCACTACAAATCATTTGAGA 1146
:|||||:
seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-461-599-16
seq_documentation_block:
: Sequence 16, Application US/08461599
: Patent No. 5658777
: GENERAL INFORMATION:
: APPLICANT: FUJII, Toshio
: APPLICANT: IWAMATSU, Akihiko
: APPLICANT: KOSHIMOTO, Hiroyuki
: APPLICANT: MINETOKI, Toshiyuki
: APPLICANT: BOGAKI, Takayuki
: APPLICANT: NAGASAWA, Naoshi
: TITLE OF INVENTION: ALCOHOL ACETYLTRANSFERASE GENES AND USE
: TITLE OF INVENTION: THEREOF
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington, D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461,599
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER: US 08/077,939
: FILING DATE: 18-JUN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 62997/1993
: FILING DATE: 26-FEB-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 184328/1992
: FILING DATE: 18-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 49441/105 KYPA
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELE: 904136
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1974 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 346..1923
: US-08-461-599-16

alignment_scores:
Quality: 78.00 Length: 208
Ratio: 0.743 Gaps: 10
Percent Similarity: 50.481 Percent Identity: 22.115

alignment_block:
US-09-528-682-1 x US-08-461-599-16 ..
Align seg 1/1 to: US-08-461-599-16 from: 1 to: 1974

68 SerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTySerThrTyTy 84
:|||||:
528 TCAGCTCACATTTAGCTTGGAGGAAATCTGCGTGAATAATCCAACTCTTT 577
84 rIleTyValIleAlaThrAla...ProAsnMetPheAsnValAsnAsp 99
:|||||:
578 TACATATTGTTCTTACACACAGATGCCCAATCATCAAAATTTTATCCG 627
100 ValLeuGlyValTySer...ProHisProTyGluGlnGluValSerAl 115
:|||||:
628 AGTTCGAATACATTTTCACGGCCACATCCAGTCATGATATATCTCAGT 677
115 a.....LeuGlyIleProTySerGlnIleTyGlyT 127
:|||||:
678 ATTCAAGAATGAAATGAGTGTGCTTCATGAAACAACCTGAGT 727
127 rPYrArGValAsnPhenGlyValIleAspGluArgLeuHisArgAsnArg 143
:|||||:
728 ACAGTCGAGTAATGAAGCAATATTGAAAGAAATTCAAAATAGTAGCGGT 777
144 GluTyArGAspArgTyTyArG.....AsnLeuAsnIle..... 155
:|||||:
828 CTTTGGACCAACAGACCGAGTGGCGGCTAATTGCTCTTCAGAGAGAC 877
778 TCCTATACAGCAAAATTTTAACTTACTACACACTTGCATTCCTTA 827
156 ....AlaProAlaGluAspGlyTyTyArGLeuAlaGlyPheProProAspH 171
:|||||:
828 CTTTGGACCAACAGACCGAGTGGCGGCTAATTGCTCTTCAGAGAGAC 877
171 ts...GlnAlaTrpArgGluGluProTrpIleHisHisAlaProGlnGly 186
:|||||:
878 ACACAGAAAAGTGAGCAAAATTTATCTTGTATCTATATCATTCAGATGCT 927
187 CysGlyAsnSerSerArgTrpTrpIleThrGlyAspTrpCysAsnGluGluH 203
:|||||:
928 GATGGTCGCTCTTCGATCCACTTTTTCATGAT...TTAAGAGAGCAAT 974
```



```

US-08-461-621-16

alignment_scores:
    Quality: 78.00      Length: 208
    Ratio: 0.743      Gaps: 10
    Percent Similarity: 50.481      Percent Identity: 22.115

alignment block:
US-09-528-682-1 x US-08-461-621-16  ..

Align seg 1/1  to: US-08-461-621-16  from: 1  to: 1974

68 SerAlaHisIleuAlaGlyInserIleLeuSerGlyTyrSerThrTyrTy 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528 TCAGCTCACATTGCTTGGAGGAAATCTGCCGAAAAATCCAACTCTT 577
84 rIleTyr.ValIleAlaThrAla...ProAsnMetPheAsnValAsnsp 99
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
578 TACATATTGTTCTTACCAACAAGATGGCCAAATCATGAAATATTATTCG 627
100 ValLeuGlyValIlyrSer...ProHisProTyrGluGlnIlyuValSerAl 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628 AGTTCGGAATACTATTTCACGCGCCACATCCAGTCGATATTATCTCAGT 677
115 a.....LeuGlyGlyIleProTyrSerGlnIleTyrGlyT 127
:|||||:|||||:|||||:|||||:|||||:|||||:
678 ATTACAAGATTTGAACCTGAGTGTGTGGTTCTTCATGACCAACCTGAGT 727
127 rPTyrArgValAsnPheGlyValIleAspGluValGluHisArgAsnArg 143
:|||||:|||||:|||||:|||||:|||||:|||||:
728 ACGATGCAGTAATGAAGCAAAATATTGAAGAATTTCAAAATAGTAAGGCT 777
144 GluTyrArgAspArgTyrTyrArg.....AsnLeuAsnIle.... 155
:|||||:|||||:|||||:|||||:|||||:|||||:
778 TCCTATACGCAAAAATTTTAAACTTACTACCACTTTGACTATTCCTTA 827
156 ...AlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAsn 171
:|||||:|||||:|||||:|||||:|||||:|||||:
828 CTTTGGACCAACGAGCCGATGGCGCGCTTAATTTGCTTCCCAAGAAAGC 877
171 is...GlnAlaThrPArgGluGluProThrIleHisAlaIleProGlnGly 186
:|||||:|||||:|||||:|||||:|||||:|||||:
878 ACGACAGAAAGTGGAGAAATTTATCTTGTATGTAATCATTCGATGCT 927
187 CysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsnGluGluTh 203
|||||:|||||:|||||:|||||:|||||:|||||:
928 GATGGTCGGCTCGATCCACTTTTTCATGAT...TTAAGACGACGAT 974
203 rGlnAsnIleSerThr.....IleTyrIleuArg 213
:|||||:|||||:|||||:|||||:|||||:|||||:
975 AATTAATATTATTAACCTCCACCAAAAAAATTAGATTATTTTCAAGTCG 1024
213 IuTyrGlnSerIlyValIlyAsrGlnIlePheSerAspTyrGlnInserGlu 229
||:|||||:|||||:|||||:|||||:|||||:|||||:
1025 AGAGAGATTACCAATTATTATGAGAACTTCAGAAACGATCGAAAGGTG 1074
230 ValASP..... 231
:|||||:|||||:|||||:|||||:|||||:|||||:
1075 ATAGACTTTTAGACCACCGTACTGTTTATTCGGAAGTCACTTCTTGGG 1124
232 ....IleTyrAsnArgIleArg 237
|||||:|||||:|||||:|||||:|||||:|||||:
1125 TTTCATCTACAAATCATTTTGGACA 1146

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-465-334-16
seq_documentation_block:
: Sequence 16, Application US/08465334
: Patent No. 5728412
: GENERAL INFORMATION:
: APPLICANT: FUJII, Toshio
: APPLICANT: IWAMATSU, Akihito

```



```

APPLICANT: YOSHIMOTO, HIROYUKI
APPLICANT: MINETOKI, TOSHIKATA
APPLICANT: BOGAKI, TAKAYUKI
APPLICANT: NAGASAWA, NAOSHI
TITLE OF INVENTION: ALCOHOL ACETYLPTRANSFERASE GENES AND USE
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,334
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,939
FILING DATE: 18-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 62997/1993
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 184328/1992
FILING DATE: 18-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/101 KYPA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1974 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 346..1923
US-08-465-334-16

alignment_scores:
      Quality: 78.00      Length: 208
      Ratio: 0.743      Gaps: 10
Percent Similarity: 50.481 Percent Identity: 22.115

alignment_block:
US-09-528-682-1 x US-08-465-334-16 ..
Align seg 1/1 to: US-08-465-334-16 from: 1 to: 1974

68 SetAlHIsleuAlaGlnSerIleLeuSerGlyTyrSerThrTyrTyr 84
|||||
528 TCAGCTCACATAGCTTTGAGGAAATCTGCGTGAATAATCCACACTCTT 577
84 rIeTyrValIleAlaThrAla...ProAsnMetPheAsnValAsnAsp 99
: |||
578 TACATATGTGTTACCAACAAGATGCGCAATATCATGAAATTTATTCGC 627
: |||
100 ValLeuGlyValTyrSer...ProHisProTyrGluGlnGluValSerAl 115
|||||
628 AGTTCGAATACATATTCACGGCCACATCCAGTCATGATATATTCACAGT 677

```

```

115 a.....LeuGlyGlyIleProTyrSerGlnIleTyrGlyT 127
: |||
678 ATTACAGAATTTGAACCTGAGTGTGCTTCATGAAACACCTGACT 727
127 rPTrArgValAsnPheGlyValIleAspGluArgLeuHisArgAsnArg 143
: |||
728 ACAGTCAGTAAATGACGAATATTTAGAGAAATTCAAAATAGTAAGGCT 777
144 GluTTrArgAspArgTyrTyrArg.....AsnLeuAsnIle..... 155
: |||
778 TCCTATACCTGCAGAAATTTTAACTTACATCCACTTGTGACTATCTCTTA 827
156 ....AlaProAlaGluAspGlyTTrArgLeuAlaGlyPheProProAsp 171
828 CTTTGGACACAGACGACGAGTGGCGGCTAATTTGCTTCCAGAAAGCG 877
171 Is...GlnAlaTrpArgGluGluProTrpIleHisAlaProGlnGly 186
|| : |||
878 ACACAGAAAGCTGAGAAATTTATCTTGTATCTATCATCTTCATGCTCT 927
187 CysGlyAsnSerSerArgTrpIleThrGlyAspThrCysAsnGluGlu 203
||| : |||
928 GATGTCGCGCTTCGATCCACTTTTTCATGAT...TTAAGACAGCAAT 974
203 rGlnAsnLeuSerThr.....IleTyrLeuArg 213
: |||
975 AATTAATATTAAACCTCCACCAAAAATTAATTAATTTCAATTTCAAGTAC 1024
213 IuTTrGlnSerIleValIleAspGlnIlePheSerAspTyrGlnSerGlu 229
|| : |||
1025 AGAGGATTTACCAATTTATGAGAACTTCACAGACGATCGAAAGGCT 1074
230 ValAsp..... 231
1075 ATAGACTTTAGACCACCGTACTGTTTATTCGAAAGTCACTCTTTCGGG 1124
232 ....IleTyrAsnArgIleArg 237
1125 TTTCATCTTACATCATCTTTGAGA 1146

seq_name: /cgn2_6/plodata/2/1na/6A_COMB.seq:us-08-990-470A-1
seq documentation block:
; Sequence 1, Application US/08990470A
; Patent No. 6123942
; GENERAL INFORMATION:
; APPLICANT: Stojiljkovic, Igor
; APPLICANT: SO, Magdalene
; APPLICANT: Hwa, Vivian
; APPLICANT: Heffron, Fred
; APPLICANT: Nassif, Xavier
; TITLE OF INVENTION: No. 6123942el Bacterial Hemoglobin Receptor
; TITLE OF INVENTION: Genes and Uses
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDowell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,470A
; FILING DATE: 15-DEC-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 6123942nan, Kevin E
; REGISTRATION NUMBER: 35,303

```



```

REFERENCE/DOCKET NUMBER: 94,784-K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3318 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 470..2845
US-08-990-470A-1

alignment_scores:
Quality: 77.50 Length: 369
Ratio: 0.601 Gaps: 15
Percent Similarity: 34.959 Percent Identity: 18.157

alignment_block:
US-09-528-682-1 x US-08-990-470A-1 ..
Align seg 1/1 to: US-08-990-470A-1 from: 1 to: 3318

7 ArgAlaSPSerArgProProAspGluIleLysArgSerGlyGlyLeuMe 23
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1601 CGTTGGACGACCAATCCG.....TTGCACCTCGGGGGGGCGACA 1641
23 tProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsnIle 40
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1642 CCGCGCTGCGTTTAAACTTTCGCCAGCGCGCGGATTTGAAAACCTAA 1691
40 snLeuTyrAspHis..... 44
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1692 ACCCGAGCATTTACTTTCAGCGCGCGTGTTCGAAACCACGAGCAGT 1741
45 .....AlaArgGlyThrGlnThrGlyPhe..... 52
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1742 ATCCAGCATCCGGTGAAACCACTACGAGTTTCTACTGTCTGACCA 1791
53 .....ValArgTyrAsp..... 56
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1792 AATTCAATGACGACGTTCTAGTACCGCGGAGGTATCCGTTACGATC 1841
56 ..... 56
1842 ATACCAAAATGACGCTCAGGAATGAATGCCAGTGCATGCTTGAC 1891
57 .....AspGlyTyrValSer 61
1892 AAAACACCGCTGACGCCACACTTATTAAGCTGGAGCGGTTTGTGCG 1941
61 tRthSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 78
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1942 CTGGCGCGCGCACTGAATCAGGCTTGCGCGTGGTTAGACACTTACT 1991
78 ergGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsn.. 93
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1992 CCGGCTAC.....CGTGTCCCAATGCG 2014
94 .....MetPheAsnValAsnAspValLeuGlyValTyrSerProH 107
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2015 TCCGAAGTGTATTCTACTTACACACGCGTTGGGTAATGCGTCCCAA 2064
107 sPro..... 108
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2065 TCCCAACCTGAAGCCGACGACGACGACCCACACCTCTCTGTGCAAG 2114
108 ..... 108

```

```

2115 GCGCAGCAGAAAAAGTACTTGGATGCCAACTGTATCAAAAGCAATFAC 2164
109 .....TyrGlnGlnGluValSerAlaLeuGlyIle.. 119
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2165 CGCAATTTCTGTCTGAAGAGCAGAAAGCTGACACGACGCGCGCATGTCA 2214
120 .....ProTyrSerG 123
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2215 CTGTACTGACATGATTTACTACTAGGTATGTAGCAATCTTATTTCG 2264
123 InIleTyrGlyTyrTyrArgValAlaAspHegIlyValIleAspGluArgLeu 139
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2265 AAAAATCGAATGCAAGATGCAAAAT.....ATCGACAGGCCACGA 2305
140 HisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAl 156
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2306 ATCCGCGCTCGAGCTGACGGGCGCT.....CTGAATGTGA 2343
156 aProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisGln 173
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2344 C.....AAAGTAGCGTCTTTGTTCTT 2365
173 LatrParGluGluProTyrIleHisHisAlaProGlnGlyCysGlyAsn 189
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2366 .....GAGGCTGGAAACTGTTCCGCTCGCTGGGTTATGCGAAA 2404
190 SerSerArgThrIleThrGlyAspThrCysAsnGlnGluThrGlnAsnLe 206
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2405 AGCANA.....CTGTGCGCGACAAACAGCTGCTGTCCACCCAGCGGT 2448
206 userThrIleTyrLeuArgGluTyrGlnSer..... 216
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2449 GAAAGTATTTCCCGGTATGCAATATGAAGTCCGAGGAAAAAATGGGGCG 2498
217 .....LysValIysArgGlnIle 222
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2499 TGTTCTCCCGCTGACCTATCTGGCGCGGAAAAAGTCAAAGACGCGCAA 2548
223 PheSerAspTyrGlnSerGlu.....ValAspIleTyrAsnArgI 236
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2549 TACACGTTTATGAACAAGAGCGCTGGGTTACGCTTTTGCAAGAAAAAG 2598
236 eArgAsp 238
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2599 AAAAGAT 2605
seq_name: /cgn2_6/ptodata/2/lna/6B.COMB.seq:US-08-817-707-1
seq_documentation_block:
; Sequence 1, Application US/08817707
; Patent No. 6277382
; GENERAL INFORMATION:
; APPLICANT: Stojiljkovic, Igor
; APPLICANT: So, Magdalene
; APPLICANT: Hwa, Vivian
; APPLICANT: Heffron, Fred
; APPLICANT: Nessif, Xavier
; TITLE OF INVENTION: No. 6277382el Bacterial Hemoglobin Receptor
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenuen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

```



```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,707
FILING DATE: 19-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6277382nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 94,784-J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3318 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 470..2845
US-08-817-707-1

Alignment_scores:
Quality: 77.50 Length: 369
Ratio: 0.601 Gaps: 15
Percent Similarity: 34.959 Percent Identity: 18.157

Alignment block:
US-09-528-682-1 x US-08-817-707-1 ..

Align seg 1/1 to: US-08-817-707-1 from: 1 to: 3318

7 ArgAlaAspSerArgProProAspGluIleLeuArgSerGlyLeuMet 23
||| |||||:|||||
1601 CGTTGGACAGCATCCG.....TTCGAACCTGGGGGGGGGCGACA 1641
23 tProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsnIle 40
::: ::::: |||||
1642 CGCCGTGCTTAAACCTTGGCGAGCGCGCGGTATTGAAACCTAA 1691
40 snLeuTyrAspHis.....ValArgTyrAsp 44
|| |||||
1692 ACCGCGCATATTAATCTCAGCGCGCGGTGTTGAAACACACAGCAGT 1741
45 .....AlaArgGlyThrGlnThrGlyPhe..... 52
|| |||||
1742 ATCCAGATCCGGTAAACACCACTACAGGTTCTCAGTGTGACCA 1791
53 .....ValArgTyrAsp 56
|| |||||
1792 AATTCAATGGAACGAGCGTTCAGTAGCGCGCAGGTATCCGTTAGATC 1841
56 ..... 56
1842 ATACCAAAATGACGCTCAGATATGAAATGCCAGTGTCACTGTGTGAC 1891
57 .....AspGlyTyrValSer 61
1892 AAAACACCCGCTGCAGCCAACTTAAAGCGTGGAGCGGTTGTGCGG 1941
61 rThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 78
::: |||||:|||||
1942 CTGGGGGGGCACTGAATCAGCTTGGCGTTCGCTTACGATTAATCT 1991
78 eGclYtyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsn... 93
|||||||
1992 CCGGCTAC.....CGTGTCCCAATGCG 2014
94 .....MetPheAsnValAsnAspValLeuGlyValTyrSerProH 107
||||: ||| |||:

```

```

2015 TCAGAGTGTATTTCACCTTACAAACCAACGTTGGGGTAATTGGCTGCCAA 2064
107 sPro..... 108
::|||
2065 TCACAACCTGAAAGCGAGCGACAGCACACCACCCCTCTCTGTGCAAG 2114
108 ..... 108
2115 GCCGAGCGAAAAAGTACTTGGATGCCAAGCTGTATCAACCAATTAC 2164
109 .....TyrGlnGlnGluValSerAlaLeuGlyLeu... 119
|||||:|||||
2165 CGCAATTCTCTGTGAAGAGCAGAAGCTACACACAGCGCGCATGTGCA 2214
120 .....ProTyrSerG 123
|||||||
2215 CTGTACTCAGATGAATTACTACTACGTATGTGTAGCAATCTTAATCCG 2264
123 InIleTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArgLeu 139
:: ||| |||||
2265 AAAAAGTGAATGCGAGATGCAAAAT.....ATGCAAGAGCCAGA 2305
140 HisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAl 156
||||: |||
2306 ATCCGCGGTCTCGAGCTGACGGCGCT.....CTGAATGTGCA 2343
156 aProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisGln 173
::: |||||: |||
2344 C.....AAAGTACGCTCTTGTGTTCT..... 2365
173 LatrPargGluGluProTyrPheHisHisAlaProGlnGlyCysGlyAsn 189
||| ||| |||
2366 .....GAGGCTGGAAACTGTTGCGCTGCGCTGCTTATGCGAAA 2404
190 SerSerArgThrIleThrGlyAspThrCysAsnGluGluThrGlnAsnLe 206
||||: |||||: |||
2405 AGCAAA.....CTGTGCGCGACACACAGCTGTGTCCACCCAGCCGT 2448
206 uSerThrIleTyrLeuArgGluTyrGlnSer..... 216
|||||: |||||
2449 GAAGATGATTCGCGTATCGATGCAATGCAAGTCCGAGCAAAAAATGGGCG 2498
217 .....LysValLysArgGlnIle 222
2499 TGTTCTCCCGCTGACCTATCTGGCGCGGAAAAAGGTCAAGACGCGCAA 2548
223 PheSerAspTyrGlnSerGlu.....ValAspIleTyrAsnArgI 236
::: |||||: |||
2549 TACACGTTTATGAAAAACAAAGGCTGGGGTACCGCTTTCAGAAAAAGGT 2598
236 eArgAsp 238
::: |||
2599 AAAAGAT 2605

seq_name: /cgn2.6/ptodata/2/lna/6A_COMB.seq:US-09-007-999-1

seq_documentation_block:
; Sequence 1, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: Latexes in Paper Manufacture
; CURRENT APPLICATION NUMBER: US/09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4460
; TYPE: DNA

```


36 GlnMetAsnIleasnLeuTyraSPHisAla.....ArgGlyThrGI 49
::: |||||::: :::: |||
3153 AATAAGAAATAACTTCCTCCTAAACATGTGTTAAACCAAGATAGTCA 3202


```

: FILE REFERENCE: 50121
: CURRENT APPLICATION NUMBER: US/09/128,450
: CURRENT FILING DATE: 1998-08-03
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 27
: LENGTH: 1322
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (101)..(865)
US-09-128-450-27

alignment_scores:
      Quality: 77.00      Length: 124
      Ratio: 1.149      Gaps: 6
Percent Similarity: 54.032      Percent Identity: 24.194

alignment_block:
US-09-528-682-1 x US-09-128-450-27 ..

Align seg 1/1 to: US-09-128-450-27 from: 1 to: 1322

114 SerAlaLeuGIyGIYleProTyrSerGlnIleTyrGIYTPYrAgya 130
      ::::::::::::::: ||| ||| ::
455 GCAGTACGTGGGGCCCTT.....GGTGCCTACATCT 486

130 LAsnPheGIYAlIleAspGIYArgLeuHISArgSmaRgIuYrAga 147
      ::::::::::: ::||| :::::::::::::::
487 GGGGAGCGCGGTGACGAGGCCCATGATTCATTYTGCAACAGCTGGGAGG 536

147 sPaRgTYrTYrAgy...AsnLeuAsnIleAlaProAlaGIuAspGIYr 162
      ||| ||| ||| ::| ||| ::|
537 ACCGCTACACCGTGAAMACATGTACCGCTACCTAACCAAGTGTACTAC 586

163 ArgLeuAlaGIYpHeProAspHisGlnAlaTrpArgGIuIuProTr 179
      ||| ||| ||| ::::::: ::
587 AGG.....CCAGTGGATCAGTACACCAACGAGAACAACTT 621

179 pIleHSHISAlaProGlnGIYCySGIYAsn.....SerSerArgT 193
      ::::::: ||| ||| :::::::
622 CGGCGCC.....GACTGGCTCAATTCACCATCAAGCGGCACA 659

193 hrIleThrGIYAspThrCysAsnGIuGIuThrGlnAspLeuSerThrIle 209
      ||::: ||| ::||| ||| :::::::
660 CGGTCCACCCACACCAAGGGGGAG.....AACTTCAACGAGACC 700

210 TYrLeuArgGIuTYrGlnSerLysValSylsArgGlnIlePheSerAspTY 226
      ::::::: :: ||| :::::::
701 GATGTCAAGATGATGAGCGCGTGTGGAGCAGATGTCGTCACCCACTA 750

226 rGlnSerGIuValAspIleTYr 233
      ||| ::||| ::| |||
751 CCAGAGGAGTCCAGGCCCTAT 772

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-823-494-27

seq_documentation_block:
: Sequence 27, Application US/09823494
: Patent No. 6355610
: GENERAL INFORMATION:
: APPLICANT: Chasebro, Bruce W
: APPLICANT: Caughey, Byron W
: APPLICANT: Chabry, Joelle
: APPLICANT: Priola, Susette
: TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prio
: TITLE OF INVENTION: Protein
: FILE REFERENCE: 50121
: CURRENT APPLICATION NUMBER: US/09/823,494
: CURRENT FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: 09/128,450

```


; PRIOR FILING DATE: 1998-08-03
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 27
 ; LENGTH: 1322
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(865)
 ; US-09-823-494-27

alignment_scores:
 Quality: 77.00 Length: 124
 Ratio: 1.149 Gaps: 6
 Percent Similarity: 54.032 Percent Identity: 24.194

alignment_block:
US-09-528-682-1 x US-09-823-494-27 ..

Align seg 1/1 to: US-09-823-494-27 from: 1 to: 1322

```

114 SerAlaIeuGlyIleProTyrSerGlnIleTyrGlyTyrTyrGva 130
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
455 GCAGTAGTGGGGGGCTT.....GTTGGCTACATGCT 486
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
130 IAsPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
487 GGGAGCGCGCTGAGCAGGCCCATGATTCATTGGCAACGACGTGGAG 536
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
147 SPARGTYTYTArg..AsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
537 ACCGCTACTACCGTAACATGATGACCGCTACCTTAACCAAGTGTACTAC 586
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
163 ArgLeuAlaGlyPheProProAspHisGlnAlaTyrArgGluTyr 179
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
587 AGG.....CCAGTGCATCAGTACAGCAACGAGAACACTT 621
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
179 pIleHisIleAlaProGlnGlyCysGlyAsn.....SerSerArgT 193
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
622 CGTGCAC.....GACTCGTCATATATCATCATCAAGCAGACACA 659
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
193 hrIleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIle 209
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
660 CGGTACACACCAACCAAGGGGAG.....AACTTACCGAGAC 700
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
210 TyrLeuArgGlyTyrGlnSerLysValLysArgGlnIlePheSerAspTyr 226
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
701 GATGTGAAGATGATGAGCGCGTGTGAGACGATGTGCTGACCCAGTA 750
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
226 rGlnSerGluValAspIleTyr 233
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
751 CCAGAGAGAGTCCAGCGCTTAT 772
    : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```

seq_name: /cgn2_6/ptodata/2/lna/5A.COMB.seq:US-08-326-670A-1

seq_documentation_block:

; Sequence 1, Application US/08326670A
 ; Patent No. 5698438
 ; GENERAL INFORMATION:
 ; APPLICANT: Stojiljkovic, Igor
 ; APPLICANT: So, Magdalene
 ; APPLICANT: Hwa, Vivian
 ; APPLICANT: Hefron, Fred
 ; APPLICANT: Nassif, Xavier
 ; TITLE OF INVENTION: A NO. 5698438el Bacterial Hemoglobin Receptor Gene and Uses
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Allegretti & Mitcoff, Ltd.
 ; STREET: 10 South Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: Illinois

; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/326.670A
 ; FILING DATE: 18 OCT 1994
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5698438han, Kevin E
 ; REGISTRATION NUMBER: 35,303
 ; REFERENCE/DOCKET NUMBER: 94,784
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; TELEX: 810-221-8317
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3318 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 470..2848
 ; US-08-326-670A-1

alignment_scores:
 Quality: 76.50 Length: 369
 Ratio: 0.593 Gaps: 15
 Percent Similarity: 34.959 Percent Identity: 18.157

alignment_block:
US-09-528-682-1 x US-08-326-670A-1 ..

Align seg 1/1 to: US-08-326-670A-1 from: 1 to: 3318

```

7 ArgAlaAspSerArgProProAspGluIleLysArgSerGlyLeuKe 23
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1601 CGTTTGGACAGCCATCGG.....TTGCACCTGGGGGGGGGAGACA 1641
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
23 tProAlaGlyHisAsnGluTyrPheAspArgGlyThrGlnMetAsnIle 40
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1642 CCGCGTGTCTTAAACCTTCCGACGCCGCGTGTATTTGAACCTAA 1691
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
40 snLeuTyrAspHis..... 44
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1692 ACCGCGACGATTTTACTTTCAGCGCGCGTGTGTGGAACCCACAGT 1741
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1742 ATCCAGCATCCGTTGAACCAACCACTACGAGTTTCACTGTGACCA 1791
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
53 .....ValArgTyrAsp. 56
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1792 AATTCAATGAGACAGGTGTTAGTAGCCGCGCAGGATTCGTTAGCAG 1841
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
56 ..... 56
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1842 ATACCAAAATGACCGCTCAGCAATTGAATGCCAGATGCTATGCTTGCAC 1891
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
57 .....AspGlyTyrValse 61
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1892 AAAACACCGCTGCAGCAACACTTATAAGCTGGAGCGGTTTTCGCG 1941
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
61 rThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 78
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
1942 CTTCGGCGGCGCAACTGATCAGCGCTTGGCGTGTGCGTTACGACATTA 1991
    : : : : : : : : : : : : : : : : : : : : : : : : : :
    : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```



```

78  erGIYrSerThrTYrTYrIleValIleAlaThrAlaProAsn... 93
    |||||
1992  CGGGCTAC.....CGTGTCCCAATGCG 2014
94  .....MetPheAsnValAsnAspValIleuGIYrSerProH 107
    |||||
2015  TCGAAGTGTATTCTACTACAAACGAGGTGGGTATTGGCGCCAA 2064
107  sPro..... 108
    |||
2065  TCCCAACCTGAAGCGGACGACAGCACACCCTCTCTGTCAAG 2114
108  ..... 108
2115  GCGGACGCAAAAAGTACTTTGGATCCCACTGTATCAAGCAATAC 2164
109  .....TYrGIuGIuValSerAlaIleuGIYrIle.. 119
    |||||
2165  CGCAATTTCTGTCTGAAGAGACAGACCTGACCAAGCGCGATGTACAG 2214
120  .....ProTYrSerG 123
    |||||
2215  CTGTACTCAGATGAAATTACTACTAGCGTATGTAGCAATCTTATTCG 2264
123  InIleTYrGIYrTYrTYrValAsnPhelGIYrIleAspGIuArgLeu 139
    |||||
2265  AAAAAGTGAATGGCAGATGCAGAAAT.....ATCGACAAAGCGCAGA 2305
140  HisArgAsnArgGIuTYrArgAspArgTYrTYrArgAsnLeuAsnIleAl 156
    |||||
2306  ATCCGGGTATTCAGAGCTGACGGGCGGT.....CTGATGTGTGA 2343
156  apPAlaGIuAspGIYrTYrArgIleuAlaGIYrPheProProAspHisGlnA 173
    |||||
2344  C.....AAAGTACGCTCTTTGTTCCT..... 2365
173  IaTYrArgGIuGIuProTYrIleHisHisAlaProGIuGIYrCysGIYrAsn 189
    |||||
2366  .....GAGGCTGGAAGCTGTTCGGCTCGCGGTATGCGAA 2404
190  SerSerArgThrIleThrGIYrAspThrCysAsnGIuGIuThrGlnAsnLe 206
    |||||
2405  AGCAAA.....CTGTGGGCGACACAGCCTGTCTCCACCACCGCCTT 2448
206  uSerThrIleTYrIleuArgGIuTYrGIuInser..... 216
    |||||
2449  GAAAGTGTATTCGCGGTATGCACTATGAAGTCCGAGCAAAAATGGGCG 2498
217  .....LysValIysArgGlnIle 222
    |||||
2499  TGTTCCTCCCGCTGACCTATCTGGGCGCGAAGGTCAAAGACGCGCAA 2548
223  PheSerAspTYrGIuInserGIu.....ValAspIleTYrAsnArgI 236
    |||||
2549  TACACCTTTATGAACAAGAGCGGTGGGTACGCTTTGCAAGAAAAGCT 2598
236  eArgAsp 238
    |||||
2599  AAAAGAT 2605

```

seq_name: /sgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-482-918-47

```

seq_documentation_block:
; Sequence 47, Application US/08482918
; Patent No. 6207417
; GENERAL INFORMATION:
; APPLICANT: zsebo, Krisztina M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Suggs, Sidney V.
; APPLICANT: Martln, Francis H.
; TITLE OF INVENTION: Stem Cell Factor
; NUMBER OF SEQUENCES: 104

```

```

CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerslein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482, 918
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/33005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 5864 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: join(565..579, 1684..1797, 2693..2755, 3351..3521,
LOCATION: 3932..4088, 4314..4397, 4778..4887, 5208..5275,
LOCATION: 5677..5713)
NAME/KEY: mat_peptide
LOCATION: join(1744..1797, 2693..2755, 3351..3521, 3932
LOCATION: 4088, 4314..4397, 4778..4887, 5208..5275, 5677
LOCATION: ..5713)
US-08-482-918-47

alignment_scores:
Quality: 76.50 Length: 214
Ratio: 0.850 Gaps: 12
Percent Similarity: 42.056 Percent Identity: 25.234

alignment_block:
US-09-528-682-1 x US-08-482-918-47 ..
Align seg 1/1 to: US-08-482-918-47 from: 1 to: 5864
11 ArgProProAspGIuIleIysArgSer..... 19
|||||
245 CGCCCTCCGATTCCTCCCTCCGCGCCCTGCTTCCTTCGCGCTACCCC 294
20 .....GlyGIYrLeuMetProArgGIYrHisAsnGIuTYrP 31
|||||
295 GGGCTCGGAAGGAGAGAGCGGTGTCCGAGACAGCGCGGGGAA.... 340
31 heAspArgGIYrThrGlnMetAsnIleAsnLeuTYrAspHisAlaArgGIY 47
|||||
341 .....CTGTATAAAGCGCGCGCGC 361
48 ThrGlnThrGIYrPheValArgTYrAspAspGIYrValSerThrSerIe 64
|||||
362 TCACACCGCGCTTCGCTCC.....CGCCTCGGCGCGAAGCT 399
64 u.....SerLeuArgSerAlaHisLeuAlaGIYrGlnSerIleuSerGIY 80
|||||
400 AGAAGCGCTGCGGAGACAGAGACAGTGAAGAGGCGCGCTCGCGGCT 449

```



```

80 yr.....SerThrTyrTyrIleTyrValIle-AlaThrAl 91
450 ACCCAATGCGTGGACTATCTGCCCGCGGTCTGCGACATATGCGGAGC 499
91 aProAsnMetPheAsnValAsnAspValIleuGlyValTyrSerPro... 106
500 TCACAGAACAGCTTAAC.....GGAGTCGCGCACACCACTGCT 534
107 .....HISProTyrGluGluGluValSerAla 115
535 TTGCTGCGATGCGACCGCTGCTTCCCTATGAAAGACACAGAGTAC 584
116 LeuGlyGlyIleProTyrSerGln..IleTyrGlyTyrTyrArgValAsn 132
585 TAGGGGGCGCGGAGGCTCCAGGCTC..... 612
132 heGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArg 148
613 .....TCCAGGAAAAATCGCGCCGCG 633
149 TyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAl 165
634 .....TCCCGCGGGGA.....GC 647
165 aclyPheProAspHisGlnAlaTyrArgGluGluProTyrIleHisH 182
648 CGCGCGTCCCTGGGACTTGCACCTGGGCGTGCAGGCTGCTCCCGCG 697
182 IsAlaProGlnGlyCysGlyAsnSerArgThrIle 194
698 GTGAGACAGAGATCGGGGAGCGCGCTGTGTG 735

```

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-224-681-47

seq_documentation_block:

Sequence 47, Application US/09224681

Patent No. 6207454

GENERAL INFORMATION:

APPLICANT: Zsebo, Krisztina M.

APPLICANT: Bosseman, Robert A.

APPLICANT: Suggs, Sidney V.

APPLICANT: Martin, Francis H.

TITLE OF INVENTION: Method for Enhancing the Efficiency of Gene

TRANSFER WITH STEM CELL FACTOR (SCF) POLYPEPTIDE

NUMBER OF SEQUENCES: 104

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/224,681

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/005,893

FILING DATE: 12-JAN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/449,653

FILING DATE: 24-MAY-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/982,255

FILING DATE: 25-NOV-1992

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/589,701
FILING DATE: 01-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/573,616
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/537,198
FILING DATE: 11-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/422,383
FILING DATE: 16-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/35199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX:
INFORMATION FOR SEQ. ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 5864 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: join(565..579, 1684..1797, 2693..2755, 3351..3521, 3932..4088, 4314..4397, 4778..4887, 5208..5275, 5677..5713)
LOCATION: 5677..5713)
NAME/KEY: mat.peptide
LOCATION: join(1744..1797, 2693..2755, 3351..3521, 3932..4088, 4314..4397, 4778..4887, 5208..5275, 5677..5713)
LOCATION: ..5713)
US-09-224-681-47

```

alignment_scores:

Quality: 76.50 Length: 214
Ratio: 0.850 Gaps: 12
Percent Similarity: 42.056 Percent Identity: 25.234

alignment_block:

US-09-528-682-1 x US-09-224-681-47 ..

Align seg 1/1 to: US-09-224-681-47 from: 1 to: 5864

```

11 ArgProAspGluIleLysArgSer..... 19
|||||.....
245 CGCCCTCCGATCTCTCCCTCCGCGCCCTCCGCTTCGCTCCGCTACCC 294
20 .....GlyGlyLeuMetProArgGlyHisAsnGluTyrP 31
|||||.....
295 GCGCTCCGAGAGGAGGAGGCGGTGCGAGACGCGCGCGGAA... 340
31 heAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGly 47
|||||.....CTGTATTAAGCGCGCGGCG 361
341 .....
48 ThrGlnArgPheValArgTyrAspAspGlyTyrValSerThrSerLe 64
|||||.....
362 TCAGCAGCGCGCTTCGCTCGC.....CGCTCGCGCGGAGACT 399
64 u.....SerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyT 80
|||||.....
400 AGAGCGCTCGCGGAGGAGGAGGAGGAGGCGCTGCGCTCGCGGCT 449
80 yr.....SerThrTyrTyrIleTyrValIle-AlaThrAl 91
|||||.....
450 ACCCAATGCGTGGACTATCTGCCCGCGGTCTGCGACATATGCGGAGC 499

```



```

132 heglyValIleaspluargleuHisArgAsnArgLufTyArgspArg 148
613 .....TCCAGCAAAATCCGCCCGG 633
149 TyrTyArgAsnLeuAsnIleAlaProAlaGluAspGlyTyArgLeuAl 165
634 .....TCCCGCGGGAA.....GC 647
165 aGlyPheProAspHisGlnAlaTPrArgGluGluProTrpIleHisH 182
648 CGGCGCTCCCTGGGACTGGGCTGGGCGTGCAGAGCTGTGCTGCCG 697
182 ISAlaProGlnGlyCysGlyAsnSerSerArgThrIle 194
698 GTGAGACAAAGAGATCGCGGGAGCGCGCGTGTGTG 735

```

```

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-08-537-361E-5

```

```

seq_documentation_block:
; Sequence 5, Application US/08537361E
; Patent No. 6121037
;
; GENERAL INFORMATION:
; APPLICANT: Stojiljkovic, Igor
; APPLICANT: So, Magdalene
; APPLICANT: Hwa, Vivian
; APPLICANT: Heffron, Fred
; APPLICANT: Naself, Xavier
; TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor
; TITLE OF INVENTION: Genes and Uses
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,361E
; FILING DATE: 02-OCT-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6121037nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 94,784-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2379 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2376
; US-08-537-361E-5

```

```

alignment_scores:
  Quality: 76.00      Length: 292
  Ratio: 0.710      Gaps: 13
  Percent Similarity: 36.644      Percent Identity: 19.521

```

```

alignment_block:
US-09-528-682-1 x US-08-537-361E-5

```

```

Align seg 1/1 to: US-08-537-361E-5 from: 1 to: 2379

```

```

42 TyrAspHis.....AlaArgGlyThrGlnThrGlyPheValArgT 44
1366 TAGCATCATACCAAAATAGCGCTCAGAAATGTAATGCCAGTGTATGC 1415
45 .....AlaArgGlyThrGlnThrGlyPheValArgT 55
1416 TTGTGACAAACACCGCTCAGCAACTTATTAAGCTGC..... 1458
55 yraspaspGlyTyValSerThrSerLeuSerLeuArgSerAlaHisLeu 71
1459 .....ACGGTTTGTGCTGGCTGGCGGCGCACTGATCAGCTTGACG 1503
72 AlaGlyGlnSerIleLeuSerGlyTySerThrTyTyIleTyValI 88
1504 GTCGTTACGACATTTACTTCGGCTAC..... 1530
88 eAlaThrAlaProAsn.....MetPheAsnValAsnAspValI 101
1531 ....CGTGTCCCAATGCTCCGAGATGATTTCACTTACACACGGTT 1576
101 euGlyValTySerProHisPro..... 108
1577 CGGCTAATTGGCTGCCAATCCCACTGAAGCCGAGCCGACGACACC 1626
108 ..... 108
1627 CACACCCCTCTCTCAGAGCCGACGAGAAAGTACTTTGGATGCCAA 1676
109 .....TyGluGlnGluValS 114
1677 CCTGTATCAACGAATTACGGAATTTCTGTCTGAAGACACAACTGA 1726
114 eAlaLeuGlyGlyLe..... 119
1727 CCACACGCGCGAGTGCAGCTGATCAGATGAATTACTTACGCTATG 1776
120 .....ProTySerGlnIleTyGlyTyTrpTyArgValAsnHeg 133
1777 TGTAGCAATCTTATTCGAAACCTGGAATGGCAGATGCAAAAT..... 1821
133 yValIleAspGluArgLeuHisArgAsnArgGluTyArgAspArgTy 150
1822 ....ATGACAAAGCCGCAATCCGCGGTATCAGCTGACGCGCGCT... 1863
150 yArgAsnLeuAsnIleAlaProAlaGluAspGlyTyArgLeuAlaGly 166
1864 .....CTGAATGTGAC.....AAATGACGCTCT 1887
167 PheProProAspHisGlnAlaTPrArgGluGluProTrpIleHisAl 183
1888 TTGTGTTCT.....GAGGCGTGGAAACTGTTCGG 1916
183 aProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCys 200
1917 CTGCTGGGTATGCGAAAGCAAA.....CTGCGGCGGACAAACAGCC 1960
200 snGluGlnThrGlnAsnLeuSerThrIleTyArgLeuArgGlnTyGlnSer 216
1961 TGCTGTCCACCCAGCCGTGAAGATGATCCGCTATCAGCTATGAAG 2010
216 ..... 216
2011 CCGAGCGAAATAGGGGCTGTCTCCCGCTGACTATCTGGCGCGAA 2060
217 .LysValTyArgGlnIlePheSerAspTyArgInSerGlu.....V 230
2061 AAAGGTCAAAGACGCGCAATACACCGTTTATGAACAAACAGGCTGGG 2110
230 alAspIleTyArgAsnArgIleArgAsp 238

```


1727 CTAAACCCCTAAATCTTTAAAT 1747

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-408-095-26

seq_documentation_block:

Sequence 6, Application US/08/408095
Patent No. 5858678
GENERAL INFORMATION:
APPLICANT: Chinadurai, Govindaswamy
TITLE OF INVENTION: APOPTOSIS-REGULATING PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,095
FILING DATE: 21-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ ID NO.: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 946 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 26..505
IS-08-408-095-26

```
alignment_scores:
  quality: 75.00
  ratio: 1.19
  percent_similarity: 41.615
  length: 161
  gaps: 6
  percent_identity: 21.118
```

```

alignment_block:
US-09-528-682-1 x US-08-408-095-26/rev ..
Align seg 1/1 to reverse of: US-08-408-095-26 from: 1 to: 946

```

```

100 ValLeuGIyValIlySerProHis..... 107
    ::::::::::::::::::::
882 CTCCTCGAATCTATTCCAGTAGACTCTTTGCCAGGGCATCAATATCA 833
    ..:::..
108 .....ProTyrGluIngluValSerAlaLeug 117
    |||
    ||| |||
832 CACAGGGGTGCACCAGCAGCATCCCCCTGCCTAACCTGCCGTGTGGCAGGG 783
    |||
117 IyGIyLleProTyrSerGlnLleTyrGlyTyrPtyrYarValAsnPhcGly 133
    ::::: |||
    ::::: |||
782 CAGAGGTGAATGGCTCT.....TCCTGGGCACAGGCCACAGTTAAC 742
    ..:::..
134 ValIIeasrgLuAtgLeuHIsArGAsnArgSltutYarg..... 146
    :::::
    :::::
741 CGGTGCACATTTGCAGAGCCATTAAGAATTCATCTTTTGAAAAAAACGAAA 692
    ..:::..
147 .....AsparTyrTYTarGAsnLeuAsnIIeAlaPrOAlaGluaspG 161
    ::::: |||
    ::::: |||
691 CTGCAGTAAAAAAACAAAAAACCTGAGTATATAAACCTCAGCAGTGT 642
    ..:::..
161 IyTYrArgLeuAlaGlyPhe..... 167
    :::::
    :::::
641 CCAGCACTATCTCGGGGTTTAATATAAAAAAGCATCATGAGAAAACAGT 592
    ..:::..
168 .....ProProAspHisGlnLatrp. 174
    |||||
    |||||
591 TAAAAAGATTACAGCAGCAGGCCGCCACCTTCAGCGGAGTGTCATGGG 542
    ..:::..
175 .....ArgGluInuProTrpIleHisAlaProGlnIyc 187
    ||| |||||
    ||| |||
541 GTGGGGCCAGCCAGCCCTTGAGCCCGGGGCCCTCACTTAGCAGCAGGT 492
    :::::
187 ys.....GlyAsnSerSerArgThrIle 194
    |||
    |||
491 GCAGGCCCGCCGTGAGCAGCGGAGCAGACAGCCACAGCAGCAGCAGCAGC 442
    ::::: |||||
    ::::: |||
195 ThrGlyAspThrCysAsnGluGluThrlGlnAsn 205
    ::::: |||||
    ::::: |||
441 GCCACGACACACTGTGTCCAGACAGCACCCAGCAGC 409

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq;US-09-182-117-1

seq_documentation_block:

Sequence 1, Application US/09182117
Patent No. 6204436
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Transgenic Plant
NUMBER OF SEQUENCES: 27
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/182,117
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8012 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-182-117-1

alignment_scores:

Quality:	74.50	Length:	102
Ratio:	1.263	Gaps:	4
Percent Similarity:	57.843	Percent Identity:	24.510

alignment_block:

US-09-528-682-1 x US-09-182-117-1/rev ..

Align seg 1/1 to reverse of: US-09-182-117-1 from: 1 to: 8012

```
74 GluSerIleLeuSerGIYrSerThrTYrIleTYrValIleAlaTh 90
||| ::||| ||| ::| ::| ::| ::| ::| ::| ::| ::|
5175 CAGTTTCTCCTGCCAGGCCAGAACTTCTTTCCAGTACCTTCCTGCCG 5126
90 rAlaPro...AsnMetPheAsnValAsnAspValIleuGIYValTYrSerP 106
::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
5125 TTTCGAATGCGCGCTTGGACATACCATCCGTAATAACGGTTCAGGCAC 5076
106 rOHISProTYrGIuGIuValSerAlaLeuGIYIleProTYrSer 122
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
5075 AGCAC.....ATCAAGAGATCGCTGATGCTATCGGTGAGCGTCGCA 5032
123 GluIleTYrGIYrPYrArgValAsnPheGIYValIleAspGIuArgLe 139
::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
5031 GAACATTAC.....ATTGACGACAGTGATCGAGCGCTCGGGT 4994
139 uHISArgAsnArgGIuTYrArgAspArgTYrTYrArgAsnLeuAsnIleA 156
::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
4993 CGAGTTTACCGGTGCTTCCGCCAG.....TGCGAAATGCCGAATAT 4950
156 IaProAlaGIuAspGIYrArgLeuAlaGIYrPheProAlaSPHisGIu 172
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
4949 TCCCGTCAGAGACGGGTATCCGGTTCGTGGCAATACTCCACATCACCCAC 4900
173 AlaTrp 174
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
4899 GCTTGG 4894
```


cgcn2_6/p.todata/2/pna/US094.COMB	seq:US-09-471-275-2419	-	88.00	146.09	109.62	276
cgcn2_6/p.todata/2/pna/US094.COMB	seq:US-09-488-722A-6884	-	88.00	146.09	109.62	27
cgcn2_6/p.todata/2/pna/US094.COMB	seq:US-09-496-916A-3388	-	88.00	146.09	109.67	27
cgcn2_6/p.todata/2/pna/US095D.COMB	seq:US-09-560-875A-3388	-	88.00	146.09	109.67	27
cgcn2_6/p.todata/2/pna/US090.COMB	seq:US-09-027-080-1	+ 88.00	133.25	568.75	9717	2
cgcn2_6/p.todata/2/pna/US092A.COMB	seq:US-60-245-222-80	-	88.00	111.65	9.1e+03	8075
cgcn2_6/p.todata/2/pna/US092B.COMB	seq:US-09-948-941-654	-	88.00	110.60	1.0e+04	895
cgcn2_6/p.todata/2/pna/US099B.COMB	seq:US-09-948-941-555	-	88.00	109.81	1.2e+04	967
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-226-176-834	-	88.00	103.13	2.7e+04	186
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-223-468-834	-	88.00	103.13	2.7e+04	186
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-313-371-834	-	88.00	103.13	2.7e+04	186
cgcn2_6/p.todata/2/pna/US6011.COMB	seq:US-60-173-464-15217	-	87.00	153.46	42.58	1
cgcn2_6/p.todata/2/pna/US6016.COMB	seq:US-60-167-217-18572	-	87.00	153.43	42.73	1
cgcn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-803-736-1478	-	87.00	106.10	1.39e+04	2436
cgcn2_6/p.todata/2/pna/US090.COMB	seq:US-09-040-990-1	+ 86.00	144.33	137.96	45	1
cgcn2_6/p.todata/2/pna/US093.COMB	seq:US-09-365-530-1	+ 86.00	184.04	0.8436	45	1
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5519	-	86.00	163.11	12.35	35
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5519	-	86.00	161.88	14.47	35
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6709	-	86.00	161.88	14.62	35
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-7068	-	86.00	161.77	14.67	35
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5214	-	86.00	161.50	15.20	41
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5214	-	86.00	161.50	15.20	41
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5178	-	86.00	161.42	15.34	41
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5178	-	86.00	161.42	15.34	41
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6714	-	86.00	161.30	15.59	41
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5527	-	86.00	161.20	15.78	42
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6131	-	86.00	161.18	15.83	42
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6131	-	86.00	161.18	15.83	42
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6176	-	86.00	161.18	15.83	42
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5447	-	86.00	160.96	16.27	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5447	-	86.00	160.96	16.27	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6123	-	86.00	160.96	16.27	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6123	-	86.00	160.96	16.27	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6862	-	86.00	160.94	16.32	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6862	-	86.00	160.94	16.32	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6535	-	86.00	160.89	16.32	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-6535	-	86.00	160.89	16.47	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5829	-	86.00	160.87	16.47	43
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5829	-	86.00	160.78	16.62	44
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5829	-	86.00	160.78	16.62	44
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5287	-	86.00	160.71	16.82	44
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5287	-	86.00	160.71	16.82	44
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5601	-	86.00	160.71	16.82	44
cgcn2_6/p.todata/2/pna/US6022.COMB	seq:US-60-220-535-5601	-	86.00	160.71	16.82	44
cgcn2_6/p.todata/2/pna/US6008.COMB	seq:US-60-082-302-686	-	86.00	156.81	27.72	64
cgcn2_6/p.todata/2/pna/US6008.COMB	seq:US-60-082-302-686	-	86.00	119.95	3.1e+03	240
cgcn2_6/p.todata/2/pna/US093A.COMB	seq:US-09-514-000-375	-	86.00	114.44	6.4e+03	413
cgcn2_6/p.todata/2/pna/US093A.COMB	seq:US-09-739-449-309	-	86.00	99.20	4.5e+04	183
cgcn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-803-110-209	-	86.00	99.20	4.5e+04	183
cgcn2_6/p.todata/2/pna/US098A.COMB	seq:US-09-528-400-110	-	85.50	159.44	19.77	1
cgcn2_6/p.todata/2/pna/US099B.COMB	seq:US-09-933-524-111331	-	85.50	159.44	19.77	1
cgcn2_6/p.todata/2/pna/US099B.COMB	seq:US-09-933-524-111331	-	85.50	159.44	19.77	1
cgcn2_6/p.todata/2/pna/US098B.COMB	seq:US-09-849-526A-111331	-	85.50	159.44	19.77	1
cgcn2_6/p.todata/2/pna/US098B.COMB	seq:US-09-849-526A-111331	-	85.50	159.44	19.77	1
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	150.14	65.21	11
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	150.14	65.21	11
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116
cgcn2_6/p.todata/2/pna/US6031.COMB	seq:US-60-312-338-311	-	85.50	149.83	67.86	116

[illegible]

[illegible]

[illegible]

/cgn2_6/ptodata/2/pna/US06036_COMB.seq	:US-60-360-207-9386 +	77.00	126.50	1.4e+03
/cgn2_6/ptodata/2/pna/US0984C_COMB.seq	:US-09-887-272A-2615 +	77.00	126.18	1.4e+03
/cgn2_6/ptodata/2/pna/US0994_COMB.seq	:US-09-404-5520-5863 +	77.00	126.12	1.4e+03
/cgn2_6/ptodata/2/pna/US06036_COMB.seq	:US-60-360-207-9371 +	77.00	125.20	1.5e+03
/cgn2_6/ptodata/2/pna/US06036_COMB.seq	:US-60-360-207-9373 +	77.00	125.12	1.5e+03
/cgn2_6/ptodata/2/pna/US0974_COMB.seq	:US-09-703-708-9385 +	77.00	124.12	1.8e+03
/cgn2_6/ptodata/2/pna/US05016_COMB.seq	:US-60-164-320-9385 +	77.00	124.12	1.8e+03
/cgn2_6/ptodata/2/pna/US05018_COMB.seq	:US-60-183-791-9385 +	77.00	124.12	1.8e+03
/cgn2_6/ptodata/2/pna/US0997A_COMB.seq	:US-09-708-427-799651 +	77.00	123.06	2.1e+03
/cgn2_6/ptodata/2/pna/US0974_COMB.seq	:US-09-708-427-799651 +	77.00	121.98	2.4e+03
/cgn2_6/ptodata/2/pna/US0937B_COMB.seq	:US-09-738-626-3388 +	77.00	119.94	3.1e+03
/cgn2_6/ptodata/2/pna/US06036_COMB.seq	:US-60-360-207-18029 +	77.00	118.60	3.7e+03
/cgn2_6/ptodata/2/pna/US0964_COMB.seq	:US-09-617-099B-2 +	77.00	117.79	4.1e+03
/cgn2_6/ptodata/2/pna/US0999B_COMB.seq	:US-09-948-941-161 -	77.00	116.89	4.6e+03
/cgn2_6/ptodata/2/pna/US0999B_COMB.seq	:US-09-948-941-161 -	77.00	116.89	4.6e+03
/cgn2_6/ptodata/2/pna/US06036_COMB.seq	:US-60-360-207-7385 +	77.00	116.29	5.0e+03
/cgn2_6/ptodata/2/pna/US0994_COMB.seq	:US-09-404-520-9384 +	77.00	115.98	5.2e+03
/cgn2_6/ptodata/2/pna/US06016_COMB.seq	:US-60-167-217-569 +	77.00	114.37	6.4e+03
/cgn2_6/ptodata/2/pna/US05017_COMB.seq	:US-60-173-464-496 +	77.00	114.10	6.6e+03
/cgn2_6/ptodata/2/pna/US0964_COMB.seq	:US-09-614-150-592 +	77.00	114.10	6.6e+03
/cgn2_6/ptodata/2/pna/US06019_COMB.seq	:US-60-191-681-592 +	77.00	114.10	6.6e+03
/cgn2_6/ptodata/2/pna/US0996B_COMB.seq	:US-09-620-329-2373 -	77.00	103.43	2.6e+04
/cgn2_6/ptodata/2/pna/US0966B_COMB.seq	:US-09-702-134-14214 -	77.00	103.43	2.6e+04
/cgn2_6/ptodata/2/pna/US0998A_COMB.seq	:US-09-815-264-81780 -	77.00	103.43	2.6e+04
/cgn2_6/ptodata/2/pna/US0999A_COMB.seq	:US-09-402-540-1284 -	77.00	102.14	3.1e+04
/cgn2_6/ptodata/2/pna/US06024_COMB.seq	:US-60-248-505-281 +	77.00	98.59	4.8e+04
/cgn2_6/ptodata/2/pna/US06024_COMB.seq	:US-60-248-505-281 +	77.00	96.43	6.4e+04
/cgn2_6/ptodata/2/pna/US06024_COMB.seq	:US-60-248-505-281 +	77.00	95.45	7.3e+04
/cgn2_6/ptodata/2/pna/US05025_COMB.seq	:US-60-258-054-2 +	77.00	93.45	7.3e+04
/cgn2_6/ptodata/2/pna/US05025_COMB.seq	:US-60-258-054-2 +	77.00	93.45	7.3e+04
/cgn2_6/ptodata/2/pna/US06026_COMB.seq	:US-60-266-852-14 -	77.00	93.50	9.3e+04
/cgn2_6/ptodata/2/pna/US0974_COMB.seq	:US-09-703-708-752 +	77.00	93.50	9.3e+04
/cgn2_6/ptodata/2/pna/US06016_COMB.seq	:US-60-164-332-752 +	77.00	93.50	9.3e+04
/cgn2_6/ptodata/2/pna/US06018_COMB.seq	:US-60-183-791-752 +	77.00	89.49	1.6e+05
/cgn2_6/ptodata/2/pna/US0995A_COMB.seq	:US-09-514-000-400 +	77.00	89.49	1.6e+05
/cgn2_6/ptodata/2/pna/US0958B_COMB.seq	:US-09-534-836-349 -	77.00	85.85	2.5e+05
/cgn2_6/ptodata/2/pna/US0997B_COMB.seq	:US-09-739-449-206 +	77.00	85.85	2.5e+05
/cgn2_6/ptodata/2/pna/US0998A_COMB.seq	:US-09-803-110-206 +	77.00	81.30	4.4e+05
/cgn2_6/ptodata/2/pna/US0995A_COMB.seq	:US-09-528-237-1348 +	77.00	81.30	4.4e+05
/cgn2_6/ptodata/2/pna/US0966C_COMB.seq	:US-09-654-617-415722 +	77.00	71.84	1.5e+06
/cgn2_6/ptodata/2/pna/US0966C_COMB.seq	:US-09-654-617-415722 +	76.50	145.24	122.31
/cgn2_6/ptodata/2/pna/US0996E_COMB.seq	:US-09-684-010-915722 +	76.50	145.24	122.31
/cgn2_6/ptodata/2/pna/US0998A_COMB.seq	:US-09-804-730-9820 -	76.50	145.24	122.31
/cgn2_6/ptodata/2/pna/US0958D_COMB.seq	:US-60-189-654-502039 +	76.50	142.84	166.31
/cgn2_6/ptodata/2/pna/US0958D_COMB.seq	:US-60-189-654-502039 +	76.50	142.84	166.31
/cgn2_6/ptodata/2/pna/US0958D_COMB.seq	:US-09-553-094-6620 +	76.50	142.84	166.31
/cgn2_6/ptodata/2/pna/US0958D_COMB.seq	:US-09-553-094-6620 +	76.50	138.69	283.02
/cgn2_6/ptodata/2/pna/US0958D_COMB.seq	:US-09-553-094-6620 +	76.50	138.69	283.02
/cgn2_6/ptodata/2/pna/US06025_COMB.seq	:US-60-263-616-362-1075 +	76.50	138.00	309.43
/cgn2_6/ptodata/2/pna/US06025_COMB.seq	:US-60-263-616-362-1075 +	76.50	137.80	317.27
/cgn2_6/ptodata/2/pna/US05019_COMB.seq	:US-09-417-507-6378 +	76.50	133.80	530.42
/cgn2_6/ptodata/2/pna/US05019_COMB.seq	:US-09-417-507-6378 +	76.50	133.80	530.42
/cgn2_6/ptodata/2/pna/US094_COMB.seq	:US-09-474-436-5791 +	76.50	137.46	331.75
/cgn2_6/ptodata/2/pna/US094_COMB.seq	:US-09-474-436-5791 +	76.50	136.94	354.43
/cgn2_6/ptodata/2/pna/US094_COMB.seq	:US-09-474-436-5791 +	76.50	136.94	354.43
/cgn2_6/ptodata/2/pna/US0998C_COMB.seq	:US-09-897-516-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US0998C_COMB.seq	:US-09-897-516-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US0998C_COMB.seq	:US-09-897-516-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/US05094_COMB.seq	:US-09-415-161-1222 -	76.50	134.10	509.82
/cgn2_6/ptodata/2/pna/				

[illegible][illegible]


```
/cgn2_6/plodata/2/pna/US096E.COMB.seq:US-09-684-016-389536 + 75.50 132.98 588.95
/cgn2_6/plodata/2/pna/US0931.COMB.seq:US-09-312-544-920 + 75.50 131.42 719.60
/cgn2_6/plodata/2/pna/US099A.COMB.seq:US-09-902-540-2570 + 75.50 130.70 789.14
/cgn2_6/plodata/2/pna/US0922.COMB.seq:US-09-252-991A-13578 + 75.50 129.98 865.21
/cgn2_6/plodata/2/pna/US097C.COMB.seq:US-09-312-544-1352 + 75.50 129.00 981.28
/cgn2_6/plodata/2/pna/US097C.COMB.seq:US-09-760-475-134 + 75.50 128.20 1.1e+03
/cgn2_6/plodata/2/pna/US094A.COMB.seq:US-09-489-039A-6203 + 75.50 127.80 1.1e+03
/cgn2_6/plodata/2/pna/US097C.COMB.seq:US-09-770-175-8640 + 75.50 127.78 1.1e+03
/cgn2_6/plodata/2/pna/US094A.COMB.seq:US-09-489-039A-5847 + 75.50 127.32 1.2e+03
/cgn2_6/plodata/2/pna/US094A.COMB.seq:US-09-404-520-7915 + 75.50 126.79 1.3e+03
/cgn2_6/plodata/2/pna/US096B.COMB.seq:US-09-620-394B-3479 + 75.50 126.78 1.3e+03
/cgn2_6/plodata/2/pna/US095D.COMB.seq:US-09-565-309A-56004 + 75.50 126.56 1.4e+03
/cgn2_6/plodata/2/pna/US098C.COMB.seq:US-09-887-576-161 + 75.50 126.26 1.4e+03
/cgn2_6/plodata/2/pna/US094A.COMB.seq:US-09-488-725A-1450 + 75.50 126.26 1.4e+03
/cgn2_6/plodata/2/pna/US096C.COMB.seq:US-09-653-150A-39 + 75.50 126.26 1.4e+03
/cgn2_6/plodata/2/pna/US096A.COMB.seq:US-09-614-150-23018 + 75.50 122.91 2.1e+03
/cgn2_6/plodata/2/pna/US0917.COMB.seq:US-09-173-464-18964 + 75.50 122.91 2.1e+03
/cgn2_6/plodata/2/pna/US0919.COMB.seq:US-09-191-637-23098 + 75.50 122.91 2.1e+03
/cgn2_6/plodata/2/pna/US0919.COMB.seq:US-09-191-681-18223 + 75.50 122.91 2.1e+03
/cgn2_6/plodata/2/pna/US0919.COMB.seq:US-09-252-991A-6998 + 75.50 122.75 2.2e+03
/cgn2_6/plodata/2/pna/US0922.COMB.seq:US-09-252-991A-13765 + 75.50 122.42 2.3e+03
/cgn2_6/plodata/2/pna/US0922.COMB.seq:US-09-252-991A-1100 + 75.50 120.54 2.9e+03
/cgn2_6/plodata/2/pna/US0922.COMB.seq:US-09-252-991A-13660 + 75.50 119.65 3.3e+03
/cgn2_6/plodata/2/pna/US0917.COMB.seq:US-09-614-150-13879 + 75.50 118.05 4.0e+03
/cgn2_6/plodata/2/pna/US0919.COMB.seq:US-09-191-637-11451 + 75.50 118.04 4.0e+03
/cgn2_6/plodata/2/pna/US0919.COMB.seq:US-09-191-681-24959 + 75.50 118.04 4.0e+03
/cgn2_6/plodata/2/pna/US0927.COMB.seq:US-09-278-258-11764 + 75.50 117.77 4.1e+03
/cgn2_6/plodata/2/pna/US0917.COMB.seq:US-09-167-217-14010 + 75.50 109.84 1.1e+04
/cgn2_6/plodata/2/pna/US0917.COMB.seq:US-09-173-464-11364 + 75.50 109.44 1.2e+04
/cgn2_6/plodata/2/pna/US096B.COMB.seq:US-09-620-392-66276 + 75.50 109.44 1.2e+04
/cgn2_6/plodata/2/pna/US097A.COMB.seq:US-09-702-134-8896 + 75.50 109.44 1.2e+04
/cgn2_6/plodata/2/pna/US098A.COMB.seq:US-09-815-264-86565 - 75.50 109.44 1.2e+04
/cgn2_6/plodata/2/pna/US0914.COMB.seq:US-09-614-150-13924 + 75.50 108.66 1.3e+04
/cgn2_6/plodata/2/pna/US096A.COMB.seq:US-09-614-150-13965 + 75.50 108.66 1.3e+04
/cgn2_6/plodata/2/pna/US0919.COMB.seq:US-09-191-681-10995 + 75.50 108.66 1.3e+04
/cgn2_6/plodata/2/pna/US0932.COMB.seq:US-09-324-185-12803 + 75.50 108.64 1.3e+04
/cgn2_6/plodata/2/pna/US0906.COMB.seq:US-09-068-138-34 + 75.50 107.19 1.6e+04
/cgn2_6/plodata/2/pna/US0906.COMB.seq:US-09-068-175-596 + 75.50 107.04 1.6e+04
/cgn2_6/plodata/2/pna/US096B.COMB.seq:US-09-620-392-64855 - 75.50 106.49 1.8e+04
/cgn2_6/plodata/2/pna/US096B.COMB.seq:US-09-620-392-56466 - 75.50 106.47 1.8e+04
/cgn2_6/plodata/2/pna/US097A.COMB.seq:US-09-702-134-8898 + 75.50 106.47 1.8e+04
/cgn2_6/plodata/2/pna/US098A.COMB.seq:US-09-815-264-81799 - 75.50 105.69 1.9e+04
/cgn2_6/plodata/2/pna/US0906.COMB.seq:US-09-068-138-353 - 75.50 104.31 2.3e+04
/cgn2_6/plodata/2/pna/US099A.COMB.seq:US-09-902-540-1117 + 75.50 97.53 5.6e+04
/cgn2_6/plodata/2/pna/US097A.COMB.seq:US-09-702-134-8242 + 75.50 97.53 5.6e+04
/cgn2_6/plodata/2/pna/US098A.COMB.seq:US-09-815-264-59997 + 75.50 97.40 5.6e+04
/cgn2_6/plodata/2/pna/US096B.COMB.seq:US-09-620-392-72515 + 75.50 95.53 7.2e+04
/cgn2_6/plodata/2/pna/US0921.COMB.seq:US-09-213-178-280 + 75.50 92.59 1.0e+05
/cgn2_6/plodata/2/pna/US097A.COMB.seq:US-09-702-134-8902 - 75.50 92.59 1.0e+05
/cgn2_6/plodata/2/pna/US098A.COMB.seq:US-09-815-264-81812 - 75.50 86.24 2.4e+05
/cgn2_6/plodata/2/pna/US0924.COMB.seq:US-09-245-228-98 + 75.50 86.24 2.4e+05
/cgn2_6/plodata/2/pna/US095B.COMB.seq:US-09-534-859-879 + 75.50 85.66 2.5e+05
/cgn2_6/plodata/2/pna/US095B.COMB.seq:US-09-534-859-647 + 75.50 85.49 2.6e+05
/cgn2_6/plodata/2/pna/US095A.COMB.seq:US-09-538-736-647 + 75.50 85.49 2.6e+05
/cgn2_6/plodata/2/pna/US099B.COMB.seq:US-09-948-941-726 + 75.50 71.15 1.6e+06
/cgn2_6/plodata/2/pna/PCUTS.COMB.seq:PCUTS-30747-8 + 75.00 162.96 12.60
/cgn2_6/plodata/2/pna/PCUTS.COMB.seq:PCUTS-30747-17 + 75.00 162.96 12.60
/cgn2_6/plodata/2/pna/PCUTS.COMB.seq:PCUTS-30747-19 + 75.00 162.96 12.60
/cgn2_6/plodata/2/pna/US094A.COMB.seq:US-09-470-124-8 + 75.00 162.96 12.60
/cgn2_6/plodata/2/pna/US094A.COMB.seq:US-09-470-124-17 + 75.00 162.96 12.60
/cgn2_6/plodata/2/pna/US094A.COMB.seq:US-09-470-124-19 + 75.00 162.96 12.60
/cgn2_6/plodata/2/pna/US0920.COMB.seq:US-09-205-419-103 + 75.00 105.75 1.9e+04
/cgn2_6/plodata/2/pna/US0920.COMB.seq:US-09-205-419-104 + 75.00 105.75 1.9e+04
/cgn2_6/plodata/2/pna/US0920.COMB.seq:US-09-205-419-105 + 75.00 105.75 1.9e+04
/cgn2_6/plodata/2/pna/US0920.COMB.seq:US-09-205-419-106 + 75.00 105.75 1.9e+04
/cgn2_6/plodata/2/pna/US0920.COMB.seq:US-09-205-419-107 + 75.00 105.75 1.9e+04
/cgn2_6/plodata/2/pna/US0920.COMB.seq:US-09-205-419-108 + 75.00 105.75 1.9e+04
/cgn2_6/plodata/2/pna/US0920.COMB.seq:US-09-205-419-109 + 75.00 105.75 1.9e+04
/cgn2_6/plodata/2/pna/US0920.COMB.seq:US-09-205-419-110 + 75.00 105.75 1.9e+04
```


/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-111 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-112 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-113 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-114 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-115 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-116 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-117 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-118 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-119 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-120 + 75.00 105.75 1.9e+04 108
/cgn2_6/pdata2/pna/US6020.COMB.seq:US-60-205-419-121 + 75.00 105.75 1.9e+04 108
seq_name: /cgn2_6/pdata2/pna/US097B.COMB.seq:US-09-724-315-4
seq_documentation_block:
; Sequence 4, Application US/09724315
; GENERAL INFORMATION:
; APPLICANT: HAINES, Joel R.
; APPLICANT: ARLINGTON, Joshua
; TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
; FILE REFERENCE: APE41
; CURRENT APPLICATION NUMBER: US/09/724,315
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 5500
; TYPE: DNA
; ORGANISM: plasmid pUV2004
US-09-724-315-4

alignment_scores:

Quality: 1297.00 Length: 240
Ratio: 5.404 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.167

alignment_block:

US-09-528-682-1 x US-09-724-315-4

Align seg 1/1 to: US-09-724-315-4 from: 1 to: 5500

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIy 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3975 AATGGCACAATATATACCGTGTGACTCTAGACCCCAATGAAATGAA 4024
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4025 ACGTCCGAGGCTTATATGCCAGAGGGCATATAGACTTCGATAGAG 4074
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4075 GAACCAATATGATATTAATCTTTATGATCGCGAGAGGACCAAAAC 4124
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4125 GCGTTTGCATATATGATGACGAGATATGTTCCACTTCTCTAGTTGAG 4174
67 gSerIleHisLeuIleGlnSerIleLeuSerGlyTyrSerThrTyrT 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4175 AAGTCTCAGTATAGAGAGAGCTATATATTCAGGATATTCACCTTACT 4224
84 yTlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4225 ATATATATATATATGACGACACCAATATGTTAATGTTAATGATGTA 4274
101 LeuGlyValIleTyrSerProHisProTyrGlnGlnIleValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4275 TTAGGCGATAGACGCCCTCACCAATATGACAGAGGTTCTGCGTTAGG 4324
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGlyVal 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4325 TGGATATACCATATCTCAGATATATGATGATGCTATCGTTAATTTTGGTG 4374
```

```
134 lIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4375 TGATGATGTAACCATATATACGTAACAGGAAATATAGAACCCGTTTAC 4424
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4425 AGAAATCTGAAATATATCTCCGACAGAGATGCTTACAGATTACAGTTT 4474
167 eProProAspHisGlnAlaIlePheArgGlnGluIleProTyrPheHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4475 CCAACCGATCCACCAAGCTTGAGAGAGAAACCTGATTCATCATGACAC 4524
184 rGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4525 CACAGGTTTGGAATTCATCATCAAGAACATTTACAGTGTACTTGTAT 4574
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4575 GAGGAGACCCAGATCTGACACCAATATATCTCAGGAAATATCAATCAAA 4624
217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAlaSerIleTyr 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4625 AGTTAAGAGGACAGATATTTTCAGACTATATCTCAGAGGTTGACATATATA 4674
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
4675 ACAGAAATTCGGATGAAATTA 4694
```

seq_name: /cgn2_6/pdata2/pna/US099B.COMB.seq:US-09-950-335A-5

seq_documentation_block:

; Sequence 5, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTIO
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 5
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-950-335A-5

alignment_scores:

Quality: 1293.00 Length: 240
Ratio: 5.410 Gaps: 0
Percent Similarity: 99.583 Percent Identity: 99.167

alignment_block:

US-09-528-682-1 x US-09-950-335A-5

Align seg 1/1 to: US-09-950-335A-5 from: 1 to: 723

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleIy 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 AATGGCACAATATATACCGTGTGACTCTAGACCCCAATGAAATGAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 ACGTCCGAGGCTTATATGCCAGAGGGCATATAGACTTCGATAGAG 100
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 GAACCTCAATGATATTAATCTTTATGATCAGCGAGAGAACCAAAAC 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```



```

151 GCCTTGTGCATATGATGACGATATGTTCCACTTCTTCTAGTTTGAG 200
67 gserAlaHISLeuAlaGlyInSerIleuSerGlyTyrSerThrTyr 84
|||||
201 AAGTGTCTACTTAGCAGACAGCTATATATATCAGGATATCCACTTACT 250
84 yrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
251 ATATATATGTATTAGGACAGACCAATATGTATTATGTATATGATGTA 300
101 LeuGlyValTyrSerProHISProTyrGluGluGluValSerAlaLeuG1 117
|||||
301 TTAGGGGTATACAGCCCTCACCATTGTGAACAGAGTTTCCGCTTAGG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPrrTyrArgValAsnPhG1yV 134
|||||
351 TGGAAATCAATATCTCAGATATATGATGATGATGCTGTTAATTTGGTG 400
134 AlIleAspGluArgLeuHISArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
401 TGAATTGATGACGATTACATGGTAACAGGAAATATAGACACCGGATTAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||
451 AGAAATCTGAATATAGCTCCGGCAGAGATGGTTACAGATTAGCAGGTTT 500
167 eProProAspHISGlnAlaTPrArgGluGluProTrrIleHISHisAlaP 184
|||||
501 CCACCCGGATTCACCAAGCTTGGAGAGAACCCCTGATTCATCAATGAC 550
184 roGlnGlyScGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
551 CACAAGGTGTGGAATTCATCAGAACATCAGACGATGATCTTATAT 600
201 G1uGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
|||||
601 GAGGAGACCCAGATCTGAGCACAATATATCTCAGGGAATATCAATCAAA 650
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
|||||
651 ACTTAGAGAGCAGATATTTTTCAGACTATCAGTCAGAGGTGACATATATA 700
234 snArgIleArgAspGluLeu 240
|||||
701 ACAGAAATTCGGGATGAAATTA 720

```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq: PCT-US99-30747-3

seq_documentation_block:

Sequence 3, Application PC/TUS9930747

GENERAL INFORMATION:

APPLICANT: Boyce Thompson Institute for Plant Research at Cor

TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in

FILE REFERENCE: 4868/85427

CURRENT APPLICATION NUMBER: PCT/US99/30747

CURRENT FILING DATE: 1999-12-22

NUMBER OF SEQ ID NOS: 67

SOFTWARE: PatentIn Ver. 2.1

Seq ID NO 3

LENGTH: 782

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: CDS

LOCATION: (3)..(782)

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: E. coli

OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize

OTHER INFORMATION: expression in plants.

PCT-US99-30747-3

alignment_scores:

Quality: 1287.00 Length: 240
Ratio: 5.362 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.333

alignment_block:

US-09-528-682-1 x PCT-US99-30747-3 ..

Align seg 1/1 to: PCT-US99-30747-3 from: 1 to: 782

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIlely 17
|||||
60 AACGGAGACAAGCTCTATAGGGCTGACTAGGCCCCCAGATAGATCA 109
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPrrPheAspArg 34
|||||
110 GAGGTGTGAGGCTCATGCTCAAGGGGAGACAAATGAGTACTTTGATAGG 159
34 lYThrGlnMetAsnIleAsnLeuTyrAspHISAlaArgGlyThrGlnThr 50
|||||
160 GAACCCAAATGAACATCAACCTCTATGACCATCTAGGGGAACCCAACT 209
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuAr 67
|||||
210 GCCTTGTGAGTATGATGATGATGATGATGTTCCACCTCCCTTAGTGTGAG 259
67 gserAlaHISLeuAlaGlyInSerIleuSerGlyTyrSerThrTyr 84
|||||
260 GTCTGTCTACTTGGCAGGACAAATCCATCCTCTCAGATATCTCCACTTACT 309
84 yrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
310 ACATATATGTGATGTCTACAGCACCAACATGTTCATATGATGATGTG 359
101 LeuGlyValTyrSerProHISProTyrGluGluGluValSerAlaLeuG1 117
|||||
360 TTGGGAGTATATGACCTCACCATATGAGCAGAGAGTGTCTGCTTGGG 409
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPrrTyrArgValAsnPhG1yV 134
|||||
410 TGGAAATCCCATCTCTCCAAATCTATGATGATGATAGGCTGAACCTTGGTG 459
134 AlIleAspGluArgLeuHISArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
460 TGAATTGATGAGAGCTCCATAGGAATAGGAGATATAGGACAGATCTAT 509
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||
510 AGGAACCTCAACATAGCTCCAGCAGAGAGATGGTTATAGTTGGCAGGTTT 559
167 eProProAspHISGlnAlaTPrArgGluGluProTrrIleHISHisAlaP 184
|||||
560 CCACACAGACCCACCAAGCCTTGGAGGAGAGGCCCTGATTCACACATGAC 609
184 roGlnGlyScGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
610 CACAGAGTTGTGAGACTCTCTCAAGACCATCAGAGTGCACACTTCAAT 659
201 G1uGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerly 217
|||||
660 GAGGAGACCCCAAACTTATGACACCATCTACCTTAGAAGTAAACCAATCCA 709
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyrA 234
|||||
710 GTGAGAGAGCAAAATCTTCTCAGACTACCAATCAGAGTGCAGATCTACA 759
234 snArgIleArgAspGluLeu 240
|||||
760 ATAGGATTAGGAATGAACTTC 779

```

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq: US-09-470-124-3

seq_documentation_block:


```

; Sequence 3, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Masou
; APPLICANT: Antzen
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; NAME/KEY: CDS
; LOCATION: (3)..(782)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-3

```

```

alignment_scores:
    Quality: 1287.00      Length: 240
    Ratio: 5.362          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 98.333

```

alignment_block:

US-09-528-682-1 x US-09-470-124-3 ..

Align seg 1/1 to: US-09-470-124-3 from: 1 to: 782

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
  |||||
60 AACGGACACAGCTCTATAGGGCTGACTAGGCCCCAGATGATGATCAA 109
  |||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  |||||
110 GAGGTCTGGAGGTCTCATGCCAAGGAGGACAAATGACTTTCATATAGG 159
  |||||
34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||||
160 GAACCCAAATGATGACATCACTCTATGACCATGCTAGGGACCCAACT 209
  |||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  |||||
210 GGGCTTGTGAGGTATGATGATGATGTCTCCACCTTACCTTGTAGCT 259
  |||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
  |||||
260 GTCTGCTCACTTGGCAGACATCATCTCTCGGATACCTCCACCTACT 309
  |||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||||
310 ACATCTATGATGATGCTACAGACACAAACATGTTCAATGTAATGATG 359
  |||||
101 LeuGlyValIleAsnProHisProTyrGlnGlnIleValSerAlaLeu 117
  |||||
360 TTGGAGATGATATACCTCCATCATGATGACAGAGGTGTCTGCTTTGG 409
  |||||
117 YrIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
  |||||
410 TGGATCCCACTCCCAATCTATGATGATGATGATGATGATGATGATG 459
  |||||
134 aIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
  |||||
460 TGATTGATGAGAGGCTCCATATGATGATGATGATGATGATGATGAT 509
  |||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
  |||||

```

```

510 AGGAACCTCAACATAGCTCCAGACAGATGTTATAGGTGGACGTTT 559
  |||||
167 eProProAspHisGlnAlaThrAlaArgGluProProPheHisAlaP 184
  |||||
560 CCCACACAGACACCAAGCTGGAGAGAGACCCCTGATCCACCATGAC 609
  |||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
  |||||
610 CACAAGTGTGAGAGCTCTCAAGACCATCAGGAGGTGACACTTGCAAT 659
  |||||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
  |||||
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
  |||||
710 GGTGAAGAGCAATCTTCTCAAGCTACCAATCAATGAGGTGACATCT 759
  |||||
234 snArgIleArgAspGluLeu 240
  |||||
760 ATGAGATTAGATGATGAACTC 779

```

seq.name: /cgn2_6/plodata/2/pna/US097B_COMB.seq:US-09-724-315-6

seq_documentation_block:

```

; Sequence 6, Application US/09724315
; GENERAL INFORMATION:
; APPLICANT: HAYNES, Joel R.
; APPLICANT: ARLINGTON, Joshua
; TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
; FILE REFERENCE: APE41
; CURRENT APPLICATION NUMBER: US/09/724,315
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 5488
; TYPE: DNA
; ORGANISM: plasmid pPV2007
US-09-724-315-6

```

```

alignment_scores:
    Quality: 1269.00      Length: 236
    Ratio: 5.400          Gaps: 0
    Percent Similarity: 99.576    Percent Identity: 98.729

```

alignment_block:

US-09-528-682-1 x US-09-724-315-6 ..

Align seg 1/1 to: US-09-724-315-6 from: 1 to: 5488

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluLeu 17
  |||||
3975 AATGGGACAAATTAATTAATCTTATGATGATGATGATGATGATGAT 4024
  |||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
  |||||
4025 ACCTTCCGAGAGCTTATGCTCCAGATGCGATGATGATGATGATGAT 4074
  |||||
34 LyrThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
  |||||
4075 GAACCTCAATGATATTAATCTTATGATGATGATGATGATGATGATG 4124
  |||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
  |||||
4125 GCGCTTGTGAGATGATGATGATGATGATGATGATGATGATGATGAT 4174
  |||||
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyr 84
  |||||
4175 AAGTCTCACTTATGACGACAGAGCTTATATATATATATATATATAT 4224
  |||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
  |||||

```



```

|||||
4225 ATATATATGTTATGACGACAGACCAATATGTTATGTTAATGATGTA 4274
101 LeuGlyValIYrSerProHISProTYrGluGlnIuValSerAlaLeuG 117
4275 TTGAGCCTTAACACCCCTCACCATATGAACAGAGAGTTCTGCGTTAGG 4324
117 YGlyIleProTYrSerGlnIleTYrGlyTYrPYrArgValAsnheGly 134
4325 TGGAAATCCATATTTCTCAGATATATGATGATCGTGTATTAATTTGGTG 4374
134 allIleAspGluArgLeuHISArgAsnArgGluTYrArgAspArgTYr 150
4375 TGATGTGATGAACGATTCATCGTAAACAGGGAATATAGACCGGTATTAC 4424
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGly 167
4425 AGAAATCTGAATATAGCTCCGCGCAGAGATGCTTACAGATTACAGGTTT 4474
167 eProProAspHISGlnAlaTrpArgGluGluProTrpIleHISAlaP 184
4475 CCCACCGGATACACCAAGCTTGAGAGAGAACCCCTGGATTTCATCATG 4524
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
4525 CACAAGGTGTGGAAATTCATCAAGAACATTTACAGGTGATCTGTAAT 4574
201 GluGluThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSer 217
4575 GAGAGAGCCAGATCTGACACCAATATATCTCAGGAAATATCAATCAA 4624
217 SValIysArgGlnIlePheSerAspTYrGlnSerGluValAspIleTYr 234
4625 AGTTAAAGCAGCATATTTTCAGACTATCAGTCAAGAGTTGACATATATA 4674
234 snArgIle 236
4675 ACAGAAAT 4682

seq_name: /cgn2_6/pdata/2/pna/US082.COMB.seq:US-08-256-003-5
seq_documentation_block:
: Sequence 5, Application us/08256003
: GENERAL INFORMATION:
: APPLICANT: Domenighini, Mario
: APPLICANT: Rappuoli, Rino
: TITLE OF INVENTION: Immunogenic Detoxified Mutants of
: TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt. Their Preparation and
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation
: STREET: 4560 Horton Street
: CITY: Emeryville
: STATE: California
: COUNTRY: USA
: ZIP: 94608-2916
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/256,003
: FILING DATE: 11-NOV-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: McClung, Barbara G.
: REGISTRATION NUMBER: 33,113
: REFERENCE/DOCKET NUMBER: 0315.001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 601-2708

```

```

: TELEFAX: (510) (655-3542)
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 711 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..711
: US-08-256-003-5

alignment_scores:
Quality: 1183.00 Length: 240
Ratio: 5.189 Gaps: 2
Percent Similarity: 95.000 Percent Identity: 93.333

alignment_block:
US-09-528-682-1 x US-08-256-003-5 ..

Align seq 1/1 to: US-08-256-003-5 from: 1 to: 711

1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIleLY 17
|||||
1 AATGGCGACGATTTATACCGTGTGACTGACTTAGACCCCGAGATGAATAA 50
17 sarGserGlyGlyLeuMetProArgGlyHISAsnGluTYrPheAspArg 34
|||||
51 ACGTTCCGGAGCTTTATGCCAGAGGT...AATGAGTACTTCGATGAG 97
34 LYrTrGlnMetAsnIleAsnLeuTYrAspHISAlaArgGlyThGlnThr 50
|||||
98 GAACCTCAATGAATATTAATCTTTATGATCACCGGAGAGGAACACAAAC 147
51 GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeuAr 67
|||||
148 GCGTTTGTGAGATATGATGACGGATATGTTCCACTTCTCTTAAGTTGAG 197
67 gSerAlaHISLeuAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYrT 84
|||||
198 AAGTGCACACTTAGCAGAGAGATATATATATATGAGATATCACTTACTA 247
84 YrIleTYrValIleAlaThrAlaProAsnMetPheAsnValAsnAspAl 100
|||||
248 TATATATCTGTTATAGCA.....AATATGTTAATGTTAATGATGTA 288
101 LeuGlyValIYrSerProHISProTYrGluGlnIuValSerAlaLeuG 117
|||||
289 ATTAGCGATACAGCCCTCACCATATGACAGAGAGTTTCTGCGTTAGG 338
117 YGlyIleProTYrSerGlnIleTYrGlyTYrPYrArgValAsnheGly 134
|||||
339 TGGAAATACATATCTCAGATATATGATGATGCTATCGTTAATTTGGTG 388
134 allIleAspGluArgLeuHISArgAsnArgGluTYrArgAspArgTYr 150
|||||
389 TGATGTGATGAACGATTCATCTGTAACAGGGAATATAGACCGGTATTAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGly 167
|||||
439 AGAAATCTGAATATAGCTCCGCGCAGAGATGCTTACAGATTACAGGTTT 488
167 eProProAspHISGlnAlaTrpArgGluGluProTrpIleHISAlaP 184
|||||
489 CCCACCGGATACACCAAGCTTGAGAGAGAACCCCTGGATTTCATCATG 538
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
539 CACAAGGTGTGGAAATTCATCAAGAACCAATCAAGAGTGATGATCTGTA 588
201 GluGluThrGlnAsnLeuSerThrIleTYrLeuArgGluTYrGlnSer 217

```



```

|||||
589 GAGGAGACCCAGAAATCGAGCAATATATCTCAGGGAATATCAATCAAA 638
217 svaIIysArgInIlePheSerAspTyrGlnSerGluValAspIleTyr 234
   |||||||
639 AGTTAAGAGCGAGATATTTTCAGACTATACGTACAGAGGTTGACATATATA 688
234 snArgIleArgAspGluLeu 240
   |||||||
689 ACAGAAATTCGGGATGAATTA 708

```

seq_name: /cgn2.6/ptodata/2/pna/US090_COMB.seq:US-09-044-696-1

```

seq_documentation_block:
: Sequence 1, Application US/09044696
: GENERAL INFORMATION:
: APPLICANT: BARCHELD, GAIL.
: APPLICANT: DEL GIUDICE, GIUSEPPE
: APPLICANT: RAPPUOLI, RINO
: TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL
: TITLE OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
: ADDRESS: R440
: STREET: P.O. BOX 8097
: CITY: EMERYVILLE
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94662-8097
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/044,696
: FILING DATE: 18-MAR-1998
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/041,227
: FILING DATE: 21-MAR-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: HARBIN, ALISA A.
: REGISTRATION NUMBER: 33,895
: REFERENCE/DOCKET NUMBER: 1393.002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 655-8730
: TELEFAX: (510) 655-3542
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 711 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..708
: US-09-044-696-1

```

alignment_scores:

```

Quality: 1183.00      Length: 240
Ratio: 5.189          Gaps: 2
Percent Similarity: 95.000      Percent Identity: 93.333

```

alignment_block:

US-09-528-682-1 x US-09-044-696-1 ..

Align seg 1/1 to: US-09-044-696-1 from: 1 to: 711

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLys 17

```

|||||
1 AATGCCGACGAGATATATACCGTGGACTAGACCCACAGATGAATAAAA 50
17 sArgSerGlyGlyLeuMetProAlaGlyHisAsnGluTyrPheAspArg 34
   |||
51 ACGTTCCGGAGCTCTTATGCCACAGAGT...AAGAGTACTTTCATAGAG 97
34 lYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
   |||||||
98 GAATCAATGAATATATATCTTATATCATCAGCGGAGAGACCAAAACC 147
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
   |||||||
148 GCGTTTGTCAAGATATGATGACGCGATATGTTTCCACTCTCTAGTTAG 197
67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTyrSerThrTyr 84
   |||||||
198 AAGTGTCACTTAGCAGACAGCATATATATATATATATATATATATCTA 247
84 yrlIeTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
   |||||||
248 TATATATCGTTATGCA.....AATATGTTATATGTTAATGATGTA 288
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
   |||||||
289 ATTAGCGTATACAGCCCTCACCCATATGAAACAGAGGTTCTGCGTTAG 338
117 yglYIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
   |||||||
339 TGGAAATCCCATATTCGATATATATGATGATGATGATGATGATGATG 388
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
   |||||||
389 TGATGATGTAACGATTTGATCATCGTACAGGGAATATACAGCGGTATAC 438
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
   |||||||
439 AGAATCTGAATATAGCTCCGCGCAGAGATGTTACGATTAGAGGTTT 488
167 eProProAspHisGlnIleAlaThrArgGluGluProTyrIleHisAla 184
   |||||||
489 CCCACGGATCCACAGCTTGAGAGAGAGAACCCGTGATTCATCATCACC 538
184 rGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
   |||||||
539 CACAGGTTGTGGAGATTTCATCAAGAACATACACAGGTATAGCTGAAT 588
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
   |||||||
589 GAGGAGACCCAGAAATCTGACACACATATATCTCAGGGAATATCAATCAA 638
217 svaIIysArgInIlePheSerAspTyrGlnSerGluValAspIleTyr 234
   |||||||
639 AGTTAAGAGCGAGATATTTTCAGACTATACGTACAGAGGTTGACATATA 688
234 snArgIleArgAspGluLeu 240
   |||||||
689 ACAGAAATTCGGGATGAATTA 708

```

seq_name: /cgn2.6/ptodata/2/pna/US082_COMB.seq:US-08-256-003-7

seq_documentation_block:

```

: Sequence 7, Application US/08256003
: GENERAL INFORMATION:
: APPLICANT: Domenighini, Mario
: APPLICANT: RAPPUOLI, RINO
: APPLICANT: PIZZA, Mariagrazia
: TITLE OF INVENTION: Immunogenic Detoxified Mutants of
: TITLE OF INVENTION: Cholera Toxin and of the Toxin B, Their Preparation and
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation

```



```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..720
; NAME/KEY: CDS
; LOCATION: 1..720
; US-09-044-696-3

```

alignment_scores:

```

Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

```

alignment_block:

```

US-09-528-682-1 x US-09-044-696-3 ..

```

```

Align seg 1/1 to: US-09-044-696-3 from: 1 to: 723

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleTyr 17
1 AATGATGTAAGTATATATCGGAGATTCCTAGACCTCGATGAAATATAA 50
17 SarGSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 GCAGTCAGAGTGCTTATGCCAAGAGACAGAGTACTTGACCGAG 100
34 IYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
101 GTACCAAAATGAATATCAACCTTATGATCATGCAAGAGAACACAGAG 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
151 GGATTTGTTAGGACGATGATGATGATGTTTCCACCTCAATTAAGTTGAG 200
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrT 84
201 AAGTCCCACTTAGTGGGTCAACATATGTCTGGTCATTTCTACTTATT 250
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATATACCACTGACCAACATGTTAACTTAATGATGATA 300
101 LeuGlyValIleTyrSerProHisProTyrGlnGluValSerAlaLeuG 117
301 TTAGGGGCATACAGTCTCATCCAGATGACACAGAGTTTCTGCTTTAGG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
351 TGGGATTCATACATCCCAATATATGATGATGATGATGATGATGATGAT 400
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
401 TGCCTGATGAACAATATCATGATATAGGGCTACAGAGATGATATATAC 450
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
451 AGTAACTTAGATATTTGCTCCAGACGACAGATGTTATGATGAGGAGG 500
167 eProProAspHisGlnAlaTyrPArgGluGluProTyrPheHisAlaP 184
501 CCGTCCGAGCATAGAGCTTGAGGAGAGAGCCGCTGATTCATCATGAGC 550
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
551 CGCGGGGTTGGGAAATGCTCCAAAGATCATGACATCAATATCTTCGAT 600
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
601 GAAAAAACCAGAAAGTCTAGGTAAATTTCTTACGAAATACCAATCTAA 650

```

```

217 sValIleArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
651 AGTAAAGACAAATATTTCCAGGCTATCATGATGATGATGATGATGAT 700
234 snArgIleArgAspGluLeu 240
701 ATAGAAATTAAGATGAAATTA 720

```

```

seq_name: /cgn2_6/plodata/2/pna/US099B_COMB.seq:US-09-950-335A-1

```

seq_documentation_block:

```

; Sequence 1, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTIO
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Vibrio cholerae
US-09-950-335A-1

```

alignment_scores:

```

Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

```

alignment_block:

```

US-09-528-682-1 x US-09-950-335A-1 ..

```

```

Align seg 1/1 to: US-09-950-335A-1 from: 1 to: 723

```

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleTyr 17
1 AATGATGTAAGTATATATCGGAGATTCCTAGACCTCGATGAAATATAA 50
17 SarGSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
51 GCAGTCAGAGTGCTTATGCCAAGAGACAGAGTACTTGACCGAG 100
34 IYThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
101 GTACCAAAATGAATATCAACCTTATGATCATGCAAGAGAACTCAGAGC 150
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
151 GGATTTGTTAGGACGATGATGATGATGTTTCCACCTCAATTAAGTTGAG 200
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrT 84
201 AAGTCCCACTTAGTGGGTCAACATATATGTCTGGTCATTTACTTATT 250
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
251 ATATATATGTTATATACCACTGACCAACATGTTTAACTTAATGATGATA 300
101 LeuGlyValIleTyrSerProHisProTyrGlnGluValSerAlaLeuG 117
301 TTAGGGGCATACAGTCTCATCCAGATGACAAAGAAAGTTTCTGCTTAGG 350
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
351 TGGGATTCATACATCCCAATATATGATGATGATGATGATGATGATGAT 400
134 alIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150

```



```

655 GAAAAACCCAAAGCTAGTGTAATAATTCCTTGACGATACCAATCTAA 704
217 STAllysaTgTgInlllePheSerAsTfYgInSeGluValAsPlleTyra 233
|||||.....|||
705 AATTAAAGACCAAAATTTTTCAGGCTATCAATCTGATATGATACACATA 754
234 snArGllleArGspJleuB 240
|||||.....|||
755 ATAGAATTAAGGATGAATTA 774

seq_name: /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-08-435-605-1

seq_documentation_block:
  sequence 1, Application US/08435605
  GENERAL INFORMATION:
    APPLICANT: Burnette, W. Neal
    APPLICANT: Kaslow, Harvey R.
    TITLE OF INVENTION: Recombinant DNA-derived
    TITLE OF INVENTION: Cholera Toxin
    NUMBER OF SEQUENCES: 32
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Amgen Inc.
      STREET: Amgen Center
      STREET: 1840 Denavilland Drive
      CITY: Thousand Oaks
      STATE: California
      COUNTRY: USA
      ZIP: 91320-1789
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette, 3.5 in., DS, 1.4 MB
      COMPUTER: Apple Macintosh
      OPERATING SYSTEM: Macintosh OS 7.0.
      SOFTWARE: Microsoft Word Version 5.0
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/435, 605
        FILING DATE: 05-MAY-1995
        CLASSIFICATION: 435
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 07/694,733
          FILING DATE: 02-MAY-1991
          INFORMATION FOR SEQ ID NO: 1:
            SEQUENCE CHARACTERISTICS:
              LENGTH: 777 base pairs
              TYPE: nucleic acid
              STRANDEDNESS: double stranded
              TOPOLOGY: circular
            US-08-435-605-1

alignment_scores:
  Quality: 1088.00      Length: 240
  Ratio: 4.814          Gaps: 0
  Percent similarity: 94.167  Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x US-08-435-605-1 ..

Align seg 1/1 to: US-08-435-605-1 from: 1 to: 777

1 ASnGlyAsPaRgLeuTYrArGAlAsPSeArRPrObRoAsPgluilely 17
|||.....|||
55 AATGATGATTAAGTATATATCGGCGAGATTCTAGACCTCTGATGAATAAA 104
17 sArGSeRgLyGlyLeuMeTPrOArGgLyNHASnGluTYrPhesPaRgG 34
|||.....|||
105 GCAGTCAGTGCTGTATATGCCAAGAGACAGAGTGAATCTTGACCCAG 154
34 lyThGInMeTAsnllleAsnleuTyRAsPHisAlaArGgLyThGInThr 50
|||||.....|||
155 GTACTCAATGATTAATCAACCTTTATGATCTGCGAAGAGAACTCAAGCG 204
51 GlyPheValArGTYrAsPaSPgLYTyrValSerThrSerleuSerleuAr 67
|||||.....|||

```



```

205 GGATTTGTAGGCAGATGATGATATGTTCCACCTCAATAGTTGAG 254
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
255 AAGGCCACCTTAGTGGTCAACATATATGTCGTGCAATCTTATTT 304
84 yTleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ATATATATGTTATAGCCACTGCACCCCAACATGTTTAACTTAATGATGTA 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 TTAGGGCATACAGTCTCATCCAGATGAGAACAGAAAGTTCTGCTTAGG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 TGGGATTCCTACTCCCAATATATGATGATGATCGATTCATTTTGGG 454
134 allLeaSPGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
455 TGCCTGATGACAACATTCATCCTATATAGGGGCTACAGATAGATATTAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
505 AGTAACTTAGATATATGCTCCAGCAGAGATGGTATGATGGCAGGTTT 554
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
555 CCCGCCGACATACAGCTTGAGGGAAGAGCCGTGATTCATCATCATCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
605 CCGCGGTTTGTGGATGCTCCAGATCATCATGATGATTAATACTTGGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
655 GAAAAACCCCAAGTCTAGGTAAATTCCTTCAGCAATACCAATCTAA 704
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 AGTTAAAGACAATATTTTCAGGCTATCATCATGTGATATGATACACATA 754
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
755 ATAGAAATTAAGATGAATTA 774
seq_name: /cgn2_6/plodata/2/pna/US094_COMB.seq:US-09-470-124-45
seq_documentation_block:
: Sequence 45, Application US/09470124
: GENERAL INFORMATION:
: APPLICANT: Mason
: TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
: FILE REFERENCE: 4866/84454
: CURRENT APPLICATION NUMBER: US/09/470,124
: NUMBER OF SEQ ID NOS: 67
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 45
: LENGTH: 777
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(777)
: OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
: OTHER INFORMATION: cholera toxin gene mutagenized to optimize
: US-09-470-124-45

```

```

alignment_scores:
  Quality: 1088.00      Length: 240
  Ratio: 4.814          Gaps: 0
  Percent Similarity: 94.167      Percent Identity: 81.667
alignment_block:
US-09-528-682-1 x US-09-470-124-45  ..
Align seq 1/1 to: US-09-470-124-45 from: 1 to: 777
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIle 17
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
55 AATGATGACAAGCTCTATAGGGCAGACTCAAGACCTCTGATGATGATCA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 GCAATCAGGTGCTCTTATGCCAAGGGGACAATCTGACTTGTGACAGG 154
34 yThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
155 GTACTCAGATGACATCAACCTTATATGACCATGCAAGGGGAATCAACT 204
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 GGATTTGTAGGCAATGATGATGATGATGTGCTCCACCTCATACTTAC 254
67 gSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTyrSerThrTyr 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
255 GTCGCCCACTTGGTGGGTCAACTATCTCTGTGTCACCTCTACTACT 304
84 yTleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ACATCTATGATGATGCCACGACCAACATGTTCAATGTGAATGATGTG 354
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 TTGGGACATACAGCCCTCAACCCAGATGAGCAGAGAGGTGTCTTGGG 404
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 TGAATCCCATACTCCCAATCTATGATGATGATGATGATGATGATGAT 454
134 allLeaSPGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
455 TGCCTGATGAGACATCTCATAGGAATAGGGGCTACAGGATGATCTAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
505 AGCAACTTGGACATTCCTCCAGCAGCAGATGATGATGATGATGATGAT 554
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
555 CCCTCAGAGCATAGGGCTTGAGGAGAGACCTTGATTCACCATCATCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
605 CACCAAGTTGTGGAATGCTCCAGGTCACAGCATGACCAACACTTGTAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
655 GAAAAACCCCAATCTTGGGTGATGATGCTCTTATGATGATGATGATGAT 704
217 sValLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
705 GGTGAAGGCAATCTTCTCAGGCTACCAATCTGACATTCATGACACCA 754
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
755 ATAGATCAAGATGAACATC 774
seq_name: /cgn2_6/plodata/2/pna/PCRUS_COMB.seq:PCR-US97-11719-12

```



```

1369 TCCTTGATACACATTACATTCGTAAATAGGGCGCTACAGAGTACGATATTAC 1418
151 ArgasnLeuasnIlealaProalagluaspGlyTyrargLeualaglyph 167
1419 ACTAAGTATGATATTGCTCCAGCAGCAGAGATGGTATGATGGCAGGTTT 1468
167 eProPaspHISglAlaIatPrargluGluProTPIleHISHisalap 184
1469 CCCTCCGCGAGCATAGAGCTTGGAGGGAGAGCCGTGGATTCATCATCCAC 1518
184 roGInglysgIyasnSerSerArgThrIlethrGlyaspThcysasn 200
1519 CCCCGGGTGTGGAGATGCTCCAGAGATCAACGATGAGTAATACTTGGCAT 1568
201 GluGluThhGlnasnLeuSerThrIleTyLeuargGluGlyGlnSerly 217
1569 GAAAAAACCACAAAGTCTAGGTGTAAAAATCTCTGCACAAATACCAATGAA 1618
217 sVallysArgGlnIlePheSeraspYrGlnSerGlnValaspIleTyra 234
1619 AATTAAAGACAAATATTTTCACAGCTATCATCTGATATTGATACACATA 1668
234 snArgIleargaspGluLeu 240
1669 ATGACATTAAAGCATGATATTA 1688

```

seq_name: /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-08-674-895-12

Quality:	1088.00	Length:	240
Ratio:	4.814	Gaps:	0
Percent Similarity:	94.167	Percent Identity:	81.667

alignment_block: MS-08-538-682-1 v DCT-HC07-11716-13

Align seg 1/1 to: PCT-US97-11719-12 from: 1 to: 1956

[illegible]

```
seq-documentation_block:
? Sequence 12, Application US/08674895
? GENERAL INFORMATION:
? APPLICANT: Biele, Jeffrey D.
? APPLICANT: Bookbinder, Louis H.
? APPLICANT: Hein, Mich B.
? TITLE OF INVENTION: IMMUNOCONTRACEPTION COMPOSITIONS
? TITLE OF INVENTION: CONTAINING SPERM ANTIGEN, AND METHODS OF USE
? NUMBER OF SEQUENCES: 19
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
? STREET: 10666 North Torrey Pines Road
? CITY: La Jolla
? STATE: California
? COUNTRY: United States
? ZIP: 92037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/674,895
? FILING DATE: 03-JUN-1996
? CLASSIFICATION: 424
? ATTORNEY/AGENT INFORMATION:
? NAME: Fitting, Thomas
? REGISTRATION NUMBER: 34,163
? REFERENCE/DOCKET NUMBER: 529.0
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 554-2937
? TELEFAX: (619) 554-6312
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1956 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? IS-08-674-895-12
```

US-08-674-895-12


```

Sequence 1, Application US/09724315
GENERAL INFORMATION:
APPLICANT: HAYNES, Joel R.
TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
FILE REFERENCE: AFP41
CURRENT APPLICATION NUMBER: US/09/724,315
CURRENT FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentl Ver. 2.1
SEQ ID NO 1
LENGTH: 5500
TYPE: DNA
ORGANISM: ppuv2002 plasmid
US-09-724-315-1

Alignment_scores:
Quality: 1088.00      Length: 240
Ratio: 4.814          Gaps: 0
Percent Similarity: 94.167      Percent Identity: 81.667

Alignment_block:
US-09-528-682-1 x US-09-724-315-1 ..

Align seg 1/1 to: US-09-724-315-1 from: 1 to: 5500

1 AsnGlyAspLeuTyArgAlaAspSerArgProProAspGluIlely 17
||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
3975 AATGATGATAAAGATTATATCGGACAGATTTCAGACCTCGTGAATAA 4024
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||:::|||||:::|||||:::|||||:::|||||:::|||||
4025 GCAGTCAGGTGGCTTATGCCAAGACGACAGAGTACTTTCACCGAG 4074
34 lyThrGlnMetAsnIleAsnLeuTyAspHisAlaArgGlyThrGlnThr 50
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4075 GATACCTAAATGAATATCAACCTTATCATATGCAAGACGAACTCAGACG 4124
51 GlyPheValArgTyArgAspArgGlyTyArgValSerThrSerLeuSerLeuAr 67
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4125 GCAATTGTTAAGCAGCATGATGGATATGTTTCCACCTCAATTACTTTAG 4174
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTySerThrTyT 84
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4175 AAGTGGCCACTTATGGGTCAACTATATATGTCGTCATCTACTTAT 4224
84 yTleIleTyValIleAlaIleThrAlaProAsnMetPheAsnValAsnAspVal 100
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4225 AATATAATATGTTAAGCCACATGCAACCAATGTTTAAAGTTAAAGATGA 4274
101 LeuGlyValTySerProHisProTyGlnGlnGluValSerAlaLeuGly 117
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4275 TTAGGGGCAATACAGTCTCTATCCAGATGAACACAGAACTTCTGCTTAGG 4324
117 yGlyIleProTySerGlnIleIleTyArgIleTyTyArgValAsnPheGlyV 134
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4325 TGGGATTCATCTACTCCCAATATATGATGGTATCGATTCATTTTGGGG 4374
134 aIleAspGluArgLeuHisArgAsnArgGluTyArgAspArgTyTyTyr 150
|||:::|||||:::|||||:::|||||:::|||||:::|||||
4375 TGCATTATGACAACTTACATCGTAATGGGGGCTACAGAGATAGTATATTC 4424
151 ArgAsnLeuAsnIleLeuProAlaGluAspGlyTyArgLeuAlaGlyP 167
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4425 AGTAACTTGAATATGCTCCACGACGACGATGATTATGATTTGGCAGGTTT 4474
167 eProProAspHisGlnAlaIleTyArgGluGluIleProTyPLeHisHisAlaP 184
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4475 CCTTCGGAGCAATAGAGCTTGGAGGGAAGCCGCTGGATTCATCATGAC 4524
184 roGlnIleTyGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
||| |||:::|||||:::|||||:::|||||:::|||||:::|||||

```

```

Sequence 1, Application US/09724315
GENERAL INFORMATION:
APPLICANT: HAYNES, Joel R.
TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
FILE REFERENCE: AFP41
CURRENT APPLICATION NUMBER: US/09/724,315
CURRENT FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentl Ver. 2.1
SEQ ID NO 1
LENGTH: 5500
TYPE: DNA
ORGANISM: ppuv2002 plasmid
US-09-724-315-1

Alignment_scores:
Quality: 1088.00      Length: 240
Ratio: 4.814          Gaps: 0
Percent Similarity: 94.167      Percent Identity: 81.667

Alignment_block:
US-09-528-682-1 x US-09-724-315-1 ..

Align seg 1/1 to: US-09-724-315-1 from: 1 to: 5500

1 AsnGlyAspLeuTyArgAlaAspSerArgProProAspGluIlely 17
||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
3975 AATGATGATAAAGATTATATCGGACAGATTTCAGACCTCGTGAATAA 4024
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||:::|||||:::|||||:::|||||:::|||||:::|||||
4025 GCAGTCAGGTGGCTTATGCCAAGACGACAGAGTACTTTCACCGAG 4074
34 lyThrGlnMetAsnIleAsnLeuTyAspHisAlaArgGlyThrGlnThr 50
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4075 GATACCTAAATGAATATCAACCTTATCATATGCAAGACGAACTCAGACG 4124
51 GlyPheValArgTyArgAspArgGlyTyArgValSerThrSerLeuSerLeuAr 67
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4125 GCAATTGTTAAGCAGCATGATGGATATGTTTCCACCTCAATTACTTTAG 4174
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTySerThrTyT 84
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4175 AAGTGGCCACTTATGGGTCAACTATATATGTCGTCATCTACTTAT 4224
84 yTleIleTyValIleAlaIleThrAlaProAsnMetPheAsnValAsnAspVal 100
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4225 AATATAATATGTTAAGCCACATGCAACCAATGTTTAAAGTTAAAGATGA 4274
101 LeuGlyValTySerProHisProTyGlnGlnGluValSerAlaLeuGly 117
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4275 TTAGGGGCAATACAGTCTCTATCCAGATGAACACAGAACTTCTGCTTAGG 4324
117 yGlyIleProTySerGlnIleIleTyArgIleTyTyArgValAsnPheGlyV 134
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4325 TGGGATTCATCTACTCCCAATATATGATGGTATCGATTCATTTTGGGG 4374
134 aIleAspGluArgLeuHisArgAsnArgGluTyArgAspArgTyTyTyr 150
|||:::|||||:::|||||:::|||||:::|||||:::|||||
4375 TGCATTATGACAACTTACATCGTAATGGGGGCTACAGAGATAGTATATTC 4424
151 ArgAsnLeuAsnIleLeuProAlaGluAspGlyTyArgLeuAlaGlyP 167
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4425 AGTAACTTGAATATGCTCCACGACGACGATGATTATGATTTGGCAGGTTT 4474
167 eProProAspHisGlnAlaIleTyArgGluGluIleProTyPLeHisHisAlaP 184
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
4475 CCTTCGGAGCAATAGAGCTTGAGAGGAAGCCGTGGATTCATCATGAC 4524
184 roGlnIleTyGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
||| |||:::|||||:::|||||:::|||||:::|||||:::|||||

```



```

4525 CCGCGGTTGGGATGCTCCAAAGATCATGATGAGTAACTTGGCAT 4574
201 GUGUthrglnasnleuserthriletyrleuarglutyrglnserly 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4575 GAAAAAACCAAGTAGGTAAATTCCTTACGAAATACCAATCTAA 4624
217 sVallyargGlnllepheserAspyrGlnSerGluValAspIleTyrA 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4625 AGTAAAGACAAATATTTCAGGCTATCAATCTGATATGATACACATA 4674
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4675 ATAGAAATTAAGATGAATTA 4694

seq_name: /cgn2_6/ptodata/2/pna/US086_COMB.seq:us-08-670-974-6

seq_documentation_block:
; Sequence 6, Application US/08670974
; GENERAL INFORMATION:
; APPLICANT: Mekalanos, John J.
; APPLICANT: Waldor, Matthew K.
; TITLE OF INVENTION: NOVEL CHOLERA VECTORS, VACCINES,
; TITLE OF INVENTION: AND METHODS FOR ANTIGEN DELIVERY IN GRAM-NEGATIVE
; TITLE OF INVENTION: BACTERIA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Clark & Ebling LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,974
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bleker-Brady, Kristina
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 00742/014001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6943 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; US-08-670-974-6

alignment_scores:
Quality: 1088.00 Length: 240
Ratio: 4.814 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.667

alignment_block:
US-09-528-682-1 x US-08-670-974-6 ..
Align seq 1/1 to: US-08-670-974-6 from: 1 to: 6943
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProAspGluIleLeu 17

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5745 AATGATGATTAAGTATATCGGACGAGATTCTAGACCTCTGATGAATAAA 5794
17 sArgSerGlyGlyLeuMetProArgGlyIleHisAsnGluTyrPheAspArg 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5795 GCAGTCAGGTGGCTTATGCGCAAGAGACAGAGTACTACTTTGACCGAG 5844
34 lYthrgImeTAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5845 GTACTCAATGAATATCAACCTTATGATCATGCAAGAGAACTCGAGAG 5894
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5895 GGATTGTGTAGGACGATGATGATATGTTCCACCTCAATATGATTGAG 5944
67 gSerAlaHisLeuAlaGlnSerIleLeuSerGlyTyrSerThrTyrTrp 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5945 AAGTGCCCACTTATGTTGGTCAAACTATATTGCTGCTGATCTACTTATT 5994
84 yIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5995 ATATATATGTTATAGCACATGACCAACATGTTTAACTTAATGATGTA 6044
101 LeuGlyValTyrSerProHisProTyrGlnGlnGluValSerAlaLeuG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6045 TTAGGGGCATACAGTCTCATCCAGATGACAGAAAGATTCTCTTAGG 6094
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAspGlyVal 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6095 TGGGATTCCTACTCCCAATATATGATGATGATCCAGTTCTTTGGGG 6144
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6145 TCGTGTATGAACAATTATCATCGTAATAGGGCTACAGAGATATGATTAC 6194
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6195 AGTAACTTATGATATTGCTCCAGACGAGAGGTTATGATGATGGCAGTTT 6244
167 eProProAspHisGlnAlaThrArgGlnGluProTyrPheHisAlaP 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6245 CCCTCCGACATAGAGCTTGAGGAGAGAGCCGTGATTCATCATGACAC 6294
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6295 CCGCGGTTGGGAAATGCTCAAGATCATGATGATTAFACTTCCGAT 6344
201 GUGUthrglnasnleuserthriletyrleuarglutyrglnserly 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6345 GAAAAAACCAAGTCTAGGTAAATTCCTTACGAAATACCAATCTAA 6394
217 sVallyargGlnllepheserAspyrGlnSerGluValAspIleTyrA 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6395 AGTAAAGACAAATATTTCAGGCTATCAATCTGATATGATACACATA 6444
234 snArgIleArgAspGluLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6445 ATAGAAATTAAGATGAATTA 6464

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-49

seq_documentation_block:
; Sequence 49, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49

```



```
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-49
```

```
alignment_scores:
    Quality: 1087.00      Length: 240
    Ratio: 4.831          Gaps: 0
    Percent Similarity: 93.750      Percent Identity: 81.667
```

alignment_block:
US-09-528-682-1 x PCT-US99-30747-49 ..

Align seg 1/1 to: PCT-US99-30747-49 from: 1 to: 777

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluIleTy 17
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
55 AATGATGACAAAGCTATATAGGAGAGCTCAAGACCTCTCATGATGATCAA 104
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheaspArg 34
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
105 GCATATCAGGAGGCTTATGCGCAAGGGGACAACTGAGTACTTTCACAGG 154
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
155 GTCTCTCATGATGACATCAACCTTTATGACATGACAGGAGGACCTCAACT 204
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
205 GGATTTGTGAGGCATGATGATGATATGTCACCTCCATGACTTGAG 254
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTyrSerThrTyrT 84
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
255 GTCTGCCCACTTGGAGGGGTCAACTATCTCTCTGTCGCTACTACTACT 304
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
305 ACATCTATGATGATGCTGACACTGACCCACATGTTCAATGTGAATGATGT 354
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
355 TTGGGAGCATACAGCCCTCACCAGATGACAGAGAGGTGCTGCTTTGGG 404
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAsnPheGly 134
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
405 TGGAAATCCCTACTCTCCAAATCTATGATGATGATGATGATGATGATG 454
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
134 AlIleAspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
455 TCGTTGATGACCACTCATAGGAATAGGGGCTACAGGGATAGGATATAC 504
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
505 AGCAACTTGGACATGCTGCTCAGCAGCAGATGTTATGATGATGATGATG 554
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
167 eProProAspHisGlnAlaTyrPargGluGluProTyrPheHisAlaP 184
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
555 CCTCTCAGCAATAGGGCTTGGAGGGGAGGCTTGGATTCACCATGAC 604
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
605 CACCAAGTTTGGAAATGCTCCAAAGGTCAAGCATGAGCAACACTGTGAT 654
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
201 GluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
655 GAAAGACCCCAATCTTTGGGTGAAAGTCTCTGATGATGATCAACTTAA 704
217 sValYsrArgGlnIlePheSerAspTyrGlnSerGlnValAlaPheTyr 234
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
705 GGTGAAGAGGCAATCTTCTCAGGCTTACCAATCTGACATTGACACCCACA 754
234 snArgIleArgAspGluLeu 240
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
755 ATAGATCAAGATGATGAATCTC 774
```

seq_name: /cgn2_6/plodata/2/pna/US094.COMB.seq:US-09-470-124-49

seq_documentation_block:

```
; Sequence 49, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; APPLICANT: Arntzen
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-49
```

```
alignment_scores:
    Quality: 1087.00      Length: 240
    Ratio: 4.831          Gaps: 0
    Percent Similarity: 93.750      Percent Identity: 81.667
```

alignment_block:
US-09-528-682-1 x US-09-470-124-49 ..

Align seg 1/1 to: US-09-470-124-49 from: 1 to: 777

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProaspGluIleTy 17
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
55 AATGATGACAAAGCTATATAGGAGAGCTCAAGACCTCTCATGATGATCAA 104
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheaspArg 34
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
105 GCATATCAGGAGGCTTATGCGCAAGGGGACAACTGAGTACTTTCACAGG 154
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
34 LyrhGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
155 GTACTCAGATGAAATCAACCTTTATGACCATGCAAGGGGAACTCAACT 204
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
205 GGATTTGTGAGGCATGATGATGATATGTCACCTCCATGACTTGAG 254
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTyrSerThrTyrT 84
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
255 GTCTGCCCACTTGGAGGGGTCAACTATCTCTCTGTCGCTACTACTACT 304
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
205 AGCAACTTGGACATGCTGCTCAGCAGCAGATGTTATGATGATGATGATG 254
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
167 eProProAspHisGlnAlaTyrPargGluGluProTyrPheHisAlaP 184
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
555 CCTCTCAGCAATAGGGCTTGGAGGGGAGGCTTGGATTCACCATGAC 604
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
605 CACCAAGTTTGGAAATGCTCCAAAGGTCAAGCATGAGCAACACTGTGAT 654
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
305 ACATCTATGATGATGCTGACACTGACCCACATGTTCAATGTGAATGATGT 354
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
   ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||
```



```
||||:|||||
355 TTGGAGCATACAGCCCTCACCCAGATGAGCAGAGCGTCTGCTTGGG 404
117 YGLIIPProtyrSerGlnIleTyrglyTrpTyraValAsnPhelgy 134
|||||
405 TGGATCCCACTACTCCAAATCTATGATGGATATAGGGGATGACCTTGGAG 454
134 allleaspluarGleuHisArgAsnArgGluTyraGAspArgTyrr 150
|||||
455 TGCCTGATGAGCACTCCATATAGAAATAGGGGCTACAGGATAGTACTAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyraArgLeuAlaGly 167
505 AGCAACTTGGACATCTGCTCCAGCAGCAGCATGCTATGATGGCAGGTTT 554
167 eProAspHisGlnAlaTrpArgGluGluProTrpIleHisAlaP 184
555 CCCCTCAGAGCATAGGGCTTGGAGGAGGAGCCCTGGATTCACCATGCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
605 CACCAAGCTTGGAAATGCTCCAGAGTCACAGCATGAGCAGCAGCTTGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyraLeuArgGluTyrglnSer 217
655 GAAAGACCCCAATCTTGGGTGAGTCACTCTGATGATGACCAATCTAA 704
217 sValIysArgGlnIlePheSerAspTyrglnSerGluValAspIleTyra 234
705 GGTAAAGAGCAATCTTCTCAGGCTACCAATCTGACATTGACACCACCA 754
234 snArgIleArgAspGluLeu 240
755 ATAGATCAAGATGAAGTCTC 774
```

```
seq_name: /cgn2_6/ptodata/2/pna/PCtUS_COMB.seq: PCT-US99-30747-47
seq_documentation_block:
; Sequence 47, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-47
```

```
alignment_scores:
Quality: 1084.00 Length: 240
Ratio: 4.796 Gaps: 0
Percent Similarity: 94.167 Percent Identity: 81.250
```

alignment_block:

US-09-528-682-1 x PCT-US99-30747-47 ..

Align seq 1/1 to: PCT-US99-30747-47 from: 1 to: 777

1 AsnGlyAspArgLeuTyraGlnAlaAspSerArgProProAspGluIleLeu 17

```
|||||
55 AATGATGACAAGCTCTATAGGAGAGACTCAAGACCTCTGATGATGATCA 104
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTrpTyraPhaAspArg 34
|||||
105 GCATTCAGGTGGTCTTATGCGAAGGGAGACATCTGCTACTTGGACAGGG 154
34 LyrThrGlnMetAsnIleAsnLeuTyraAspHisAlaArgGlyThrGlnThr 50
155 GTACTCGATGATGAACATCAACCTTATGACCATGCAAGGGGAATCTCAACT 204
51 GlyPheValArgTyraAspAspGlyTyraValSerThrSerLeuSerLeuAr 67
205 GCATTTGTGAGCATGATGATGATATGTTGTCACCAAGATGATGCTTGAG 254
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrrSerThrTyrr 84
255 GTCTGCCCACTTGGTGGTAAACTATACCTCTCTGGCAGCTTCACTTACT 304
84 YrIleTyraValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
305 ACATCTATGATGATTCGCACGACGACCCACATGTTCAATGATGATGATG 354
101 LeuGlyValIyrSerProHisProTyrglnGlnGluValSerAlaLeuG 117
355 TTGGAGCATATACAGCCCTCAGCAGATGAGCAAGAGGTGCTCTCTTGGG 404
117 YGLIIPProtyrSerGlnIleTyrglyTrpTyraValAsnPhelgy 134
405 TGGATCCCACTACTCCAAATCTATGATGGATATAGGGGCTCAGCTTGGAG 454
134 allleaspluarGleuHisArgAsnArgGluTyraGAspArgTyrr 150
455 TGCCTGATGAGCAACTCCATATAGAAATAGGGGCTACAGGATAGTACTAC 504
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyraArgLeuAlaGly 167
505 AGCAACTTGGACATCTGCTCCAGCAGCAGCATGCTATGATGGCAGGTTT 554
167 eProAspHisGlnAlaTrpArgGluGluProTrpIleHisAlaP 184
555 CCCCTCAGAGCATAGGGCTTGGAGGAGGAGCCCTTGATTCACCATGCAC 604
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
605 CACCAAGCTTGGAAATGCTCCAGAGTCACAGCATGAGCAGCAGCTTGAT 654
201 GluGluThrGlnAsnLeuSerThrIleTyraLeuArgGluTyrglnSer 217
655 GAAAGACCCCAATCTTGGGTGAGTCACTCTGATGATGACCAATCTAA 704
217 sValIysArgGlnIlePheSerAspTyrglnSerGluValAspIleTyra 234
705 GGTAAAGAGCAATCTTCTCAGGCTACCAATCTGACATTGACACCACCA 754
234 snArgIleArgAspGluLeu 240
755 ATAGATCAAGATGAAGTCTC 774
```

```
seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq: US-09-470-124-47
seq_documentation_block:
; Sequence 47, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; APPLICANT: Arintzen
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
```



```

; SEQ ID NO 47
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-47

```

```

alignment_scores:
  Quality: 1084.00      Length: 240
  Ratio: 4.796         Gaps: 0
  Percent Similarity: 94.167   Percent Identity: 81.250

```

```

alignment_block:
US-09-528-682-1 x US-09-470-124-47

```

```

Align seg 1/1 to: US-09-470-124-47 from: 1 to: 777

```

```

1  AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIlely 17
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
55  AATGATGACAGGCTCTATAGGCGAGACTCAAGACCTCTGATGATGATCA 104
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17  sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
105  GCATATCAGGTGGTCTTATGCCAAGGGGACATCTGACTTGTGACAGG 154
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34  LythrGlnMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrGlnThr 50
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155  GTACTCTGATGACATCAACCTTTATGACATGCAAGGGGAGACTCAACT 204
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51  GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeu 67
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
205  GGATTTGTGAGGATGATGATGATGATGATGATGATGATGATGATGATG 254
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67  gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYr 84
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255  GTCTGCCACACTGGTGGTCAAACTATCTCTGTCACCTCTACTACTACT 304
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84  YrIleTYrValIleAlaThrAlaProAsnMetPheAsnValAspAspVal 100
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
305  ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 354
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101  LeuGlyValTYrSerProHisProTYrGluGlnGluValSerAlaLeuG 117
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
355  TTGGGACATACAGCCCTCAACCCGATGACAGAGAGGTCTGCTTGGG 404
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117  yGlyIleProTYrSerGlnIleTYrGlyTYrTYrArgValAsnPheGly 134
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
405  TGGATATCCCACTACTCCCAAACTATGATGATGATGATGATGATGATG 454
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134  AlIleAspGluArgLeuHisArgAsnArgGlyTYrTYrArgAspArgTY 150
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
455  TGCCTGTGAGCAGACTCCATAGGAGGAGGCTACAGGAGATGATGATGAT 504
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151  ArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArgLeuAlaGly 167
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
505  AGCAACTTTGGACATGCTCCAGCAGAGATGATGATGATGATGATGATG 554
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167  eProProAspHisGlnAlaThrArgGluGluProTYrIleHisAlaP 184
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
555  CCCCTCCAGACATAGGCTTGGAGGAGGAGCCCTTGGATTCACCATG 604
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184  roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAs 200
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
605  CACCAAGGTTGTGGAAATGCTCCAAAGGTCAGACATGACATCACTTGT 654
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201  GlnGluThrGlnAsnLeuSerThrIleTYrLeuArgGlyTYrGlnSer 217

```

```

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
655  GAAAGACCCAACTTGGGTGAGAGTCTCTGATGATGATGATGATGATG 704
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
217  sValTYrArgGlnIlePheSerAspTYrGlnSerGluValAspIleTYr 234
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
705  GGTGAAGAGCCAAATCTTCTAGGCTTACCAATCTGACATTTGACACCA 754
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234  snArgIleArgAspGluLeu 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
755  ATAGATCAAGATGAATC 774

```

```

seq_name: /cgn2_6/plodata/2/pna/PCUTUS.COMB.seq: PCT-US99-30747-51

```

```

seq_documentation_block:

```

```

; Sequence 51, Application PC/TUS9930747

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Boyce Thompson Institute for Plant Research at Cor

```

```

; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In

```

```

; FILE REFERENCE: 4868/85427

```

```

; CURRENT APPLICATION NUMBER: PCT/US99/30747

```

```

; NUMBER OF SEQ ID NOS: 67

```

```

; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 51

```

```

; LENGTH: 777

```

```

; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence

```

```

; FEATURE:

```

```

; NAME/KEY: CDS

```

```

; LOCATION: (1)..(777)

```

```

; FEATURE:

```

```

; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae

```

```

; OTHER INFORMATION: cholera toxin gene mutagenized to optimize

```

```

; OTHER INFORMATION: expression in plants.

```

```

PCT-US99-30747-51

```

```

alignment_scores:

```

```

  Quality: 1081.00      Length: 240

```

```

  Ratio: 4.804         Gaps: 0

```

```

  Percent Similarity: 93.750   Percent Identity: 81.250

```

```

alignment_block:

```

```

US-09-528-682-1 x PCT-US99-30747-51

```

```

Align seg 1/1 to: PCT-US99-30747-51 from: 1 to: 777

```

```

1  AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIlely 17
  ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
55  AATGATGACAGGCTCTATAGGCGAGACTCAAGACCTCTGATGATGATCA 104
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17  sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
105  GCATATCAGGTGGTCTTATGCCAAGGGGACATCTGACTTGTGACAGG 154
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34  LythrGlnMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrGlnThr 50
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155  GTACTCTGATGACATCAACCTTTATGACATGCAAGGGGAGACTCAACT 204
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51  GlyPheValArgTYrAspAspGlyTYrValSerThrSerLeuSerLeu 67
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
205  GGATTTGTGAGGATGATGATGATGATGATGATGATGATGATGATGATG 254
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67  gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYr 84
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255  GTCTGCCACACTGGTGGTCAAACTATCTCTGTCACCTCTACTACTACT 304
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84  YrIleTYrValIleAlaThrAlaProAsnMetPheAsnValAspAspVal 100
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
305  ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 354
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101  LeuGlyValTYrSerProHisProTYrGluGlnGluValSerAlaLeuG 117

```


[illegible]

1 sequence, Application US/09/724315
 2 sequence, Application US/09/724315
 3 GENERAL INFORMATION:
 4 APPLICANT: HAYNES, Joel R.
 5 APPLICANT: ARRINGTON, Joshua
 6 TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
 7 FILE REFERENCE: APE41
 8 CURRENT APPLICATION NUMBER: US/09/724,315
 9 CURRENT FILING DATE: 2000-11-27
 10 NUMBER OF SEQ ID NOS: 22
 11 SOFTWARE: PatentIn Ver. 2.1


```

|||||.....
722 GGATTTGTTAGGACAGATGATGATGTTTCCACCTCAATTAAGTTGAG 771
67 gserAlahisLeuAlaGlnSerIleLeuSerGlyTyrSerThrTyr 84
|||||.....
772 AAGTGGCCACTTACTGGGTCAAACTATATGTCGTGCTACTTACTTAT 821
84 yrlIeYValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
822 ATATATATGTTATAGCCACTGCACCCAACTGTTTAACTTAATGATGA. 870
101 LeuGlyValTyrSerProHisProTyrGlnGlnGlnValSerAlaLeuG 117
|||||.....
871 TTAGGGCATACAGTCTCATCCAGATGAACAAGATTCCTGCTTAGG 920
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhelGly 134
921 TGGGATTCACATCTCCAAATATATGATGATGATGATGATGATGATG 970
134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||.....
971 TCGTTGATGAAACATTAATCATGTAATAGGGCTACAGAGATAGATATTAC 1020
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||.....
1021 AGTAACTTAGATATTGCTCCAGACAGATGTTATGATGATGATGATG 1070
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisHisAla 184
|||||.....
1071 CCCCGCCGAGATAGACCTTGAGAGGAGAGACCCGTGATTCATCATGAC 1120
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||.....
1121 CGCGGGCTTGAGGATGCTCCAGATCATGATGATGATGATGATGATG 1170
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||.....
1171 GAAATAACCCAAAGCTAGGTGTAATAATCCTTACGAAATACCAATCTAA 1220
217 svaAlaHisGlnIlePheSerAspTyrGlnSerGlyValAspIleTyr 234
|||||.....
1221 AGTTAAAGAACAAATTTTCAGCGCTATCATGATATGATGATGATG 1270
234 snArgIleArgAspGluLeu 240
|||||.....
1271 ATAGATTTAAGATGAATTA 1290
seq_name: /cgn2_6/ptodata/2/pna/PCtus_COMB.seq: PCR-US01-08582-2
seq_documentation_block:
? Sequence 2, Application PC/TUS0108582
? GENERAL INFORMATION:
? APPLICANT: UAB Research Foundation
? TITLE OF INVENTION: Chimeric Nontoxic Mutants of
? Humoral Immunity
? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hendricks and Assoc
? STREET: P.O. Box 2509
? CITY: Fairfax
? STATE: VA
? COUNTRY: US
? ZIP: 22031
? COMPUTER READABLE FORM:
? MEDIUM TYPE: floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US01/08582
? FILING DATE: 16-Mar-2001
? CLASSIFICATION: <Unknown>

```

```

? ATTORNEY/AGENT INFORMATION:
? NAME: Hendricks, Glenna
? REGISTRATION NUMBER: 32,535
? REFERENCE/DOCKET NUMBER: MCG-01
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703/425-8405
? TELEFAX: 703/425-8406
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2022 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? SEQUENCE DESCRIPTION: SEQ ID NO: 2:
PCR-US01-08582-2

alignment_scores:
Quality: 1049.00 Length: 240
Ratio: 4.683 Gaps: 0
Percent Similarity: 93.33 Percent Identity: 80.417

alignment_block:
US-09-528-682-1 x PCR-US01-08582-2 ..
Align seg 1/1 to: PCR-US01-08582-2 from: 1 to: 2022

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIle 17
||| |||.....
572 AATATGATAGTAAGTTATATCGGGCAGATTCAGACCTCGATGAAATAAA 621
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||.....
622 GCACTCAGGTGGTCTTATGCCAAGACCACAGAGTACTTGGACCGAG 671
34 lyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||.....
672 GTACTCAATAGAAATATACACCTTATGATCATGCAGAGAGACTCAGAG 721
672 GlnAlaValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
|||||.....
722 GGATTTGTTAGGCAGATGATGATGATGTTCCACCTCAATTAAGTTGAG 771
67 gserAlahisLeuAlaGlnSerIleLeuSerGlyTyrSerThrTyr 84
|||||.....
772 AAGTGGCCACTTACTGGGTCAAACTATATGTCGTGCTACTTACTTAT 821
84 yrlIeYValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
822 ATATATATGTTATAGCCACTGCACCCAACTGTTTAACTTAATGATGA. 870
101 LeuGlyValTyrSerProHisProTyrGlnGlnGlnValSerAlaLeuG 117
|||||.....
871 TTAGGGCATACAGTCTCATCCAGATGAACAAGATTCCTGCTTAGG 920
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhelGly 134
921 TGGGATTCACATCTCCAAATATATGATGATGATGATGATGATGATG 970
134 allLeaspGluArgLeuHisArgAsnArgGluTyrArgAspArgTyr 150
|||||.....
971 TCGTTGATGAAACATTAATCATGTAATAGGGCTACAGAGATAGATATTAC 1020
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
|||||.....
1021 AGTAACTTAGATATTGCTCCAGACAGAGATGTTATGATGATGATGATG 1070
167 eProProAspHisGlnAlaTrpArgGluGluProTyrIleHisHisAla 184
|||||.....
1071 CCCCGCCGAGATAGACCTTGAGAGGAGAGCCGTGATTCATCATGAC 1120

```



```

36 lmetasnleasnleutyrsphsialaargglythrghnrglyphe 52
|||||
109 AATGATATCAACCTTATGATCATGCAGAGAACTCAGCGGATTT 158
|||||
53 ValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAl 69
|||||
159 GTTAGGACGATGATGATGATGTTTCCACCTCAATTAGTTGAGAGATGC 208
|||||
69 ahlsleuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrTyrIle 86
|||||
209 CCACCTTAGTGGTCAACTATATTGTGTGTCATCTACTATTATATATAT 258
|||||
86 yvalAlaIleAlaTrpAlaProAsnMetPheAsnValAsnAspValLeuGly 102
|||||
259 ATGTTATAGCAGCAGCCAGCCACATGTTTAACTATATGATATATGGG 308
|||||
103 ValTyrSerProHisProTyrGluGlnGluValSerAlaLeuGlyIle 119
|||||
309 GCATACAGTCCCTCATCCAGATGAACAAGATTCTCTTAGTGCGGAT 358
|||||
119 eProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhcGlyValIleAla 136
|||||
359 TCCATACTCCCAATATATGATGATGATGATGATGATGATGATGATG 408
|||||
136 spgluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
409 ATGACATATACATCTAATAGGGGCTACAGATGATGATGATGATGATG 458
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhePro 169
|||||
459 TTGATATCTCTCCAGCAGCAGATGATGATGATGATGATGATGATG 508
|||||
169 oAspHisGlnAlaTrpArgGluGluProTyrIleHisIleAlaProGln 186
|||||
509 GGAGCATAGAGCTTGGAGGAGAGCGGTGATTCATCATGACCGCGG 558
|||||
186 lYcysGlyAsnSerSerArgThr 193
|||||
559 GTTGTGGAGATGCTCCAGATCA 581
|||||
seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-402-100-1

seq_documentation_block:
; Sequence 1, Application US/09402100
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Helioth
; FILE REFERENCE: 0136/06140
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: ()..()
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-1

```

```

alignment_scores:
  Quality: 252.00      Length: 112
  Ratio: 3.111        Gaps: 3
  Percent Similarity: 72.321      Percent Identity: 47.321

alignment_block:
US-09-528-682-1 x US-09-402-100-1 ..

Align seg 1/1 to: US-09-402-100-1 from: 1 to: 2385

137 GluArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArg.... 151
|||||
1557 CAGAAATATCATCTAAAGAAAGACATGCAATTCACAGCAGACCGCTCAC 1606
|||||
152 .....AsnLeuAsnIleAlaProAlaGluAspGlyTyrAla 163
|||||
1607 TTGAAGTCAATCTGAAACTTA.....CCATGCTTCCGTGGATGGC 1647
|||||
163 rgluAlaGlyPheProProAspHisGlnAla.TrpArgGluGluProTyr 179
|||||
1648 AAGCACTACTCTAAACCGCAATTAAGTGAATTCGAAAGAGCGCGT 1697
|||||
179 rIleHisIleAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrG 196
|||||
1698 GATTCATCATGACACCGCGGCTGTGGGAATGCTCCAGATCATCATCA 1747
|||||
196 lYasprHisCysAsnGluGluTyrGlnAsnLeuSerThrIleTyrLeuArg 212
|||||
1748 GTATACTCTGCGATGAAAAAACCAGTAAGTGTAGGTAAATTCCTTGAC 1797
|||||
213 GluTyrGlnSerYsValysArgGlnIlePheSerAspTyrGlnSerG 229
|||||
1798 GATTAACCAATCTAAGTTAAAGACAAATATTTTCAGGCGATCAATCTGA 1847
|||||
229 uValAspIleTyrAsnArgIleArgAspGluLeu 240
|||||
1848 TATTGATACACATATATAGATTAAGATGATTA 1881
|||||

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-402-100-3

seq_documentation_block:
; Sequence 3, Application US/09402100
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Hel
; FILE REFERENCE: 0136/06140
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 4149
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: ()..()
; OTHER INFORMATION: Recombinant DNA
US-09-402-100-3

alignment_scores:

```


Quality: 245.00 Length: 65
Ratio: 4.298 Gaps: 0
Percent Similarity: 87.692 Percent Identity: 66.154

alignment_block:

US-09-528-682-1 x US-09-402-100-3 ..

Align seg 1/1 to: US-09-402-100-3 from: 1 to: 4149

```

176 GUGUURPOTPIEHSIALAPROGLNGLYSGLYSNERSEAR 192
|||||
3451 GAAAGCCGCGATTGATCATGACCCCGGTGTGGAACTCTCCAG 3500
|||||
192 gthrlthrclyasprhrcysaenglugluthrlnasleuserthr 209
|||||
3501 ATCTTCGATCAGTAACTCTGCGATGAAAAACCAAGTCTAGGTGTA 3550
|||||
209 leryleuargglutyrlnserlysalysarglnlpheserasp 225
|||||
3551 AATTCCTTGACGATACCAATCTAAAGTTAAAGCAAAATTTTCAGGC 3600
|||||
226 Tyrlnsergluvalaspilleyrarnarglleargaspgluleu 240
|||||
3601 TATCAATCTGATATGATACACATATAGATTAAGATGATTA 3645
|||||
seq_name: /cgn2_6/prodata/2/pna/US090_COMB.seq:US-09-051-315-1

```

seq_documentation_block:

```

; Sequence 1, Application US/09051315
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co., Ltd
; APPLICANT: Kim, Byung-O
; APPLICANT: Lee, Byung-Kwang
; APPLICANT: Ioon, Suk-Won
; APPLICANT: Park, Seung-Kook
; APPLICANT: Yu, Young-Hyo
; TITLE OF INVENTION: A RECOMBINANT MICROORGANISM EXPRESSING
; TITLE OF INVENTION: AN ANTIGENIC PROTEIN, ADHESIN
; FILE REFERENCE: 0136/06164
; CURRENT APPLICATION NUMBER: US/09/051,315
; EARLIER FILING DATE: 1998-04-03
; EARLIER APPLICATION NUMBER: PCT/KR 97/00091
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence of a fusion gene prepared by ligating adhesin gene c
; OTHER INFORMATION: A2 and B subunit genes of Vibrio cholerae toxin.
; US-09-051-315-1

```

alignment_scores:

Quality: 241.50 Length: 263
Ratio: 1.963 Gaps: 6
Percent Similarity: 46.768 Percent Identity: 26.996

alignment_block:

US-09-528-682-1 x US-09-051-315-1 ..

Align seg 1/1 to: US-09-051-315-1 from: 1 to: 1516

```

1 Asnglyasparglutyrargalaspserarproproaspglulileu 17
|||||
403 AATGCGCAATGTGTTTACGCCGATCTAAAGACCAATACAGAAAA 452
|||||
17 SARSGerlyglyleumet..... 23
|||||
453 ATCAGAACCCGCGTTATATCTCCACCGCTTGGACAAAAATGGAAGG 502
|||||

```

```

24 .....Proargly.....His 27
|||||
503 TTTTAATCCCGCGTGGGTTTATTAAAGTTACCATCTAGACCTATGAGT 552
|||||
28 Asnglutyrpneaspargglythrclmecasnlleasnleutyrsaphl 44
|||||
553 GCGCAATCTTGGATCTTTTACGATGATTTGACGCGGTGGACATYCA 602
|||||
44 salargglythrclnhrgrlypnevalargtyraspaspglytyrval 61
|||||
603 AGAAATAATCTTAAACACCATTCATCAAGCATAGCGGGGTGTAGTA 652
|||||
61 erthrserleuserleuargseralshlsleualaglglinserrleu 77
|||||
653 GCAC.....ATGCTTAAGGAAAGGATATCTTAATGACGATCAG 696
|||||
78 Serglytyrserthrtyrtyrlylethyvallealathralaprasme 94
|||||
697 AGCGTTTGAATTAAGATTTTGCATAATATC.....AT 728
|||||
94 tpeasnvalasnpsvalleuglyvaltyrserprohlsprthyrglug 111
|||||
729 GCAAGAAATGACAAAAACCTCACTCAAAAGATTTGAAATCTTATCAA 778
|||||
111 lngluvalseralaleuglylyleprotyrserglnlleytyrlytrp 127
|||||
779 AAGACGCCAAGAAATTAAGGC..... 801
|||||
128 tyrrargvalasnphcglyvalilleasppluarleuhsargasnargl 144
|||||
802 .....AAAAAACCAGCA 815
|||||
144 utyrargaspargtyrtyrargasnleuasnillealaproalaclasp 161
|||||
816 ATTC..... 819
|||||
161 lytyrargleualaglypheproproasphlslnalatrpargluglu 177
|||||
820 .....CAG 822
|||||
178 prothrlehsialaproglnglycsglyasnserargtyrhl 194
|||||
823 CCGTGATTCATCATGACCCCGGTGTGGATGCTCCAGATCATC 872
|||||
194 ethrglyasprhrcysaenglugluthrlnasleuserthrlyrly 211
|||||
873 GATCAGTAACTTGCATGAAAAAACCAAGCTAGTAAATTC 922
|||||
211 euargglutyrlnserlysalysarglnlpheseraspptyrsln 227
|||||
923 TTGACGAATACCAATCTAAAGTTAAAGACAAATTTTTCAGGCTATCAA 972
|||||
228 sergluvalaspilleyrarnarglleargaspgluleu 240
|||||
973 TCTGATATGATACACATATAGATTAAGATGATTA 1011
|||||
seq_name: /cgn2_6/prodata/2/pna/US094_COMB.seq:US-09-423-493-1

```

seq_documentation_block:

```

; Sequence 1, Application US/09423493
; GENERAL INFORMATION:
; APPLICANT: DAEMONG PHARMACEUTICAL CO., LTD
; APPLICANT: Kim, B.O., et al.
; TITLE OF INVENTION: A preventive and therapeutic vaccine for
; TITLE OF INVENTION: Helicobacter pylori-associated diseases.
; FILE REFERENCE: 0136/06320
; CURRENT APPLICATION NUMBER: US/09/423,493
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: PCT/KR98/00072
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1

```



```

; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion oligonucleotide between Helicobacter pylori
; OTHER INFORMATION: and Vibrio cholerae
US-09-423-493-1

```

alignment_scores:

```

Quality: 241.50 Length: 263
Ratio: 1.963 Gaps: 6
Percent Similarity: 46.768 Percent Identity: 26.996

```

alignment_block:

```

US-09-528-682-1 x US-09-423-493-1 ..

```

```

Align seg 1/1 to: US-09-423-493-1 from: 1 to: 1516

```

```

1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIlely 17
|||||: : : : : : : : : : : : : : : : : : : : : : : : : :
403 AATGGCGAAATGTTTACGCCCGCATCTAAAGACATACAGAAAAA 452
17 sArgSerGlyGlyLeuMet..... 23
453 ATCAGAACCCGGGTTATTATCTCCACCGGTTTGACAAAATGGAAGCGG 502
24 .....ProArgGly.....His 27
503 TTTTAATCCCGCGGTGGTTATTAGGTACCATCTAGACCTATAGT 552
28 AsnGluTYrPheAspArgGlyTYrGlnMetAsnIleAsnLeuTYrAspHi 44
: : : : : : : : : : : : : : : : : : : : : : : : : :
553 GGGGAATCTTGGATCTTTTACGATGATTTGACGAGTTGACATTCGA 602
44 sAlaArgGlyThrGlnThrGlyPheValArgTYrAspAspGlyTYrValS 61
: : : : : : : : : : : : : : : : : : : : : : : : : :
603 AGAAAAATCTTAAACACACCATTCACAGCCATAGCGGGGTTAGTTA 652
61 eArThrSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleleu 77
|||||: : : : : : : : : : : : : : : : : : : : : : : : : :
653 GCACT.....ATGGTTAAGGAACGACGATATTTCTAATGACGCAATCAAG 696
78 SerGlyTYrSerThrTYrTYrIleTYrValIleAlaThrAlaProAsnMe 94
|||||: : : : : : : : : : : : : : : : : : : : : : : : : :
697 AGCCCTTGTGATATAGATTGTCAAATATC.....AT 728
94 tPheAsnValAsnAspValleuGlyValTYrSerProHisProTYrGluG 111
| : : : : : : : : : : : : : : : : : : : : : : : : : :
729 GCAAGAAATGACAAAAAACTCCTCAAAAGAAATTAGAAATCTTATCAAA 778
111 lngIuValSerAlaLeuGlyIleProTYrSerGlnIleTYrGlyTYrP 127
: : : : : : : : : : : : : : : : : : : : : : : : : :
779 AAGACGCCAAAGATTTAAAGC..... 801
128 TYrArgValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgG 144
|||||: : : : : : : : : : : : : : : : : : : : : : : : : :
802 .....AAAAGAAACCGAGA 815
144 uTYrArgAspArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAsp 161
| : : : : : : : : : : : : : : : : : : : : : : : : : :
816 ATTC..... 819
161 lTYrArgLeuAlaGlyPheProProAspHisGlnAlaTrpArgGluGlu 177
|||||: : : : : : : : : : : : : : : : : : : : : : : : : :
820 .....GAG 822
178 ProTYrIleHisAlaProGlnGlyCysGlyAsnSerSerArgTYrHi 194
|||||: : : : : : : : : : : : : : : : : : : : : : : : : :
823 CCGTGATTCATCATGCACCGCGGTTTGGAATCTCCACATCATC 872
194 eThrGlyAspThrCysAsnGluGluTYrGlnAsnLeuSerThrIleTYrL 211
: : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

873 GATCAGTAATACCTTCGATGAAAAAACCCAAAGTCTAGTGTAAATTC 922
211 euArgGluTYrGlnSerLYsVallysArgGlnIlePheSerAspTYrGln 227
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
923 TTGACGAATGCCAATCTTAAAGTTAAAGACAAATATTTTACGGCTATCAA 972
228 SerGluValAspIleTYrAsnArgIleArgAspGluLeu 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
973 TCTGATATTCATACACATATATAGAAATTAAGCATGAATTA 1011
seq_name: /cgn2.6/plodata/2/pna/US097C_COMB.seq:US-09-760-234-1

```

seq_documentation_block:

```

; Sequence 1, Application US/09760234
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co., Ltd
; APPLICANT: Kim, Byung-O
; APPLICANT: Lee, Byung-Kwang
; APPLICANT: Yoon, Suk-Won
; APPLICANT: Park, Seung-Kook
; APPLICANT: Yu, Young-Hyo
; TITLE OF INVENTION: A RECOMBINANT MICROORGANISM EXPRESSING
; TITLE OF INVENTION: AN ANTIGENIC PROTEIN, ADHESIN
; FILE REFERENCE: 0136/IE164-US1
; CURRENT APPLICATION NUMBER: US/09/760,234
; PRIOR FILING DATE: 2001-01-11
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: PCT/KR 97/00091
; PRIOR FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence of a fusion gene prepared by ligating
; OTHER INFORMATION: adhesin gene of H. pylori and A2 and B subunit genes of
; OTHER INFORMATION: Vibrio cholerae toxin
US-09-760-234-1

```

alignment_scores:

```

Quality: 241.50 Length: 263
Ratio: 1.963 Gaps: 6
Percent Similarity: 46.768 Percent Identity: 26.996

```

alignment_block:

```

US-09-528-682-1 x US-09-760-234-1 ..

```

```

Align seg 1/1 to: US-09-760-234-1 from: 1 to: 1516

```

```

1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIlely 17
|||||: : : : : : : : : : : : : : : : : : : : : : : : : :
403 AATGGCGAAATGTTTACGCCCGCATCTAAAGACATACAGAAAAA 452
17 sArgSerGlyGlyLeuMet..... 23
453 ATCAGAACCCGGGTTATTATCTCCACCGGTTTGACAAAATGGAAGCGG 502
24 .....ProArgGly.....His 27
503 TTTTAATCCCGCGGTGGTTATTAGGTACCATCTAGACCTATAGT 552
28 AsnGluTYrPheAspArgGlyTYrGlnMetAsnIleAsnLeuTYrAspHi 44
: : : : : : : : : : : : : : : : : : : : : : : : : :
553 GGGGAATCTTGGATCTTTTACGATGATTTGACGAGTTGACATTCGA 602
44 sAlaArgGlyThrGlnThrGlyPheValArgTYrAspAspGlyTYrValS 61
: : : : : : : : : : : : : : : : : : : : : : : : : :
603 AGAAAAATCTTAAACACACCATTCACAGCCATAGCGGGGTTAGTTA 652

```



```

61 eRThSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 77
   |||||
653 GCACCT.....ATGGTTAAGGAAACGATATCTTAATGACGGATCAAG 696
   |||||
78 SerGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMe 94
   |||||
697 AGCGCTTTGAAATGAATTTTGCAGAAATATC.....AT 728
   |||||
94 tPheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGluG 111
   |||||
729 GCAGAAATAGACAAAAAACTCACTCAAAAGAAATTAGAACTTATCTCAA 778
   |||||
111 InGluValSerAlaLeuGlyIleProTyrSerGlnIleTyrGlyTyr 127
   |||||
779 AAGACCCCAAGAAATTAAAGGC..... 801
   |||||
128 TyrArgValAsnDheGlyValIleAspGluArgLeuHisArgAsnArg 144
   |||||
802 .....AAGAAACCGAGA 815
   |||||
144 uTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAsp 161
   |||||
816 ATTC..... 819
   |||||
161 LyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTyrArgGluGlu 177
   |||||
820 .....GAG 822
   |||||
178 ProTyrIleHisAlaProGlnGlyCysGlyAsnSerSerArgThrIle 194
   |||||
823 CCCTGGATTCATCATGACCCGCCGGTGTGGGAATGCTCCAAAGATCATC 872
   |||||
194 eThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 211
   |||||
873 GATCAGTAATACTTGGATGAAAAAACCCAAAGCTAGGTAAATTC 922
   |||||
211 euArgGluTyrGlnSerLysValLysArgGlnIlePheSerAspTyrGln 227
   |||||
923 TTGACGAATTCATCAATCAAAAGTTAAAGCAAAATATTTTACGGCTATCAA 972
   |||||
228 SerGluValAspIleTyrAsnArgIleArgAspGluLeu 240
   |||||
973 TCTGATATTGATACATATATAGAAATTAAGATGAATTA 1011
   |||||
seq_name: /cgn2_6/ptodata/2/pna/US081.COMB.seq:US-08-133-438-3
seq_documentation_block:
; Sequence 3, Application US/08133438
; GENERAL INFORMATION:
; APPLICANT: KAPER, James B.
; APPLICANT: LEVINE, Myron M.
; TITLE OF INVENTION: VIBRIO CHOLERAE NON-01 SEROGROUP VACCINE
; TITLE OF INVENTION: STRAINS, METHODS OF MAKING SAME AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & McKenzie
; STREET: Suite 1100, 815 Connecticut Ave., N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20006-4078
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,438
; FILING DATE: 08-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/472,276
; FILING DATE: 04-MAR-1983

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/581,406
; FILING DATE: 17-FEB-1984
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/867,633
; FILING DATE: 27-MAY-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/363,383
; FILING DATE: 05-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/533,315
; FILING DATE: 05-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/821,872
; FILING DATE: 16-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/931,943
; FILING DATE: 12-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kile, Bradford E.
; REGISTRATION NUMBER: 25,223
; REFERENCE/DOCKET NUMBER: BAMCZ 0019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 452-7000
; TELEFAX: (202) 452-7074
; TELEX: 89552
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7076 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Vibrio cholerae
; STRAIN: CVD110
; US-08-133-438-3

alignment_scores:
Quality: 160.50 Length: 50
Ratio: 3.567 Gaps: 1
Percent Similarity: 90.000 Percent Identity: 64.000

alignment_block:
US-09-528-682-1 x US-08-133-438-3/rev ..
Align seg 1/1 to reverse of: US-08-133-438-3 from: 1 to: 7076
191 SerArgThrIleThrGlyAspThrCysAsnGluGluThrGlnAsnLeu 207
   |||||
1339 TCTAAGCCATCACTGATATGATACATATAGAAATTAAGATGAATTA 1293
   |||||
207 rThrIleTyrLeuArgGluTyrGlnSerLysValLysArgGlnIlePhe 224
   |||||
1292 TGTAAATTCCTTGACGAATCAATCTAAAGTTAAAGCAAAATATTTT 1243
   |||||
224 eRAspTyrGlnSerGluValAspIleTyrAsnArgIleArgAspGluLeu 240
   |||||
1242 CAGCGTATCATATGATATGATACATATAGAAATTAAGATGAATTA 1193
   |||||
seq_name: /cgn2_6/ptodata/2/pna/US081.COMB.seq:US-08-133-439-3
seq_documentation_block:
; Sequence 3, Application US/08133439
; GENERAL INFORMATION:
; APPLICANT: KAPER, James B.
; APPLICANT: LEVINE, Myron M.
; TITLE OF INVENTION: VIBRIO CHOLERAE CVD111, METHOD OF MAKING
; TITLE OF INVENTION: SAME, AND VACCINE DERIVED THEREFROM
; NUMBER OF SEQUENCES: 3

```


CORRESPONDENCE ADDRESS:
ADDRESSEE: Baker & McKenzie
STREET: Suite 1100, 815 Connecticut Ave., N.W.
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20006-4078
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/133,439
FILING DATE: 08-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/472,276
FILING DATE: 04-MAR-1983
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/581,406
FILING DATE: 17-FEB-1984
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/867,633
FILING DATE: 27-MAY-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/363,383
FILING DATE: 05-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/533,315
FILING DATE: 05-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/821,872
FILING DATE: 16-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/931,943
FILING DATE: 12-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kile, Bradford E.
REGISTRATION NUMBER: 25,223
REFERENCE/DOCKET NUMBER: BANCZ 0020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 452-7000
TELEFAX: (202) 452-7074
TELEX: 89552
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 7076 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYDROTHERICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Vibrio cholerae
STRAIN: CVD110
US-08-133-439-3

alignment_scores:
Quality: 160.50 Length: 50
Ratio: 3.567 Gaps: 1
Percent Similarity: 90.000 Percent Identity: 64.000

alignment_block:
US-09-528-682-1 x US-08-133-439-3/rev ..
Align seg 1/1 to reverse of: US-08-133-439-3 from: 1 to: 7076

191 SerArgThrIleArgLysPheCysAsnGluGluThrGlnAsnLeu 207
|||||:|||||: :|||||:|||||:|||||:|||||:|||||:
1339 TCTAGAGCGATGAGT...ATACTGCGATGAGAAAAACCCAAAGTCTAGG 1293

207 rThrIleTyrlLeuArgLysGlnSerLysValLysArgGlnIlePheS 224
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1292 TCTAAATTCCTTGACGATACCAATCTAAAGTTAAAGACAAATTTT 1243
224 eRAspTyrlGlnSerGluValAspIleTyrlAsnArgIleArgAspGluLeu 240
|| |||||:|||||: |||||:|||||: |||||:|||||: |||||:|||||:
1242 CAGGCTATCATCTGATATTGATACACATATAGATTAAAGATGAATTA 1193

seq_name: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-771-536B-15

seq_documentation_block:
; Sequence 15, Application US/09771536B
; GENERAL INFORMATION:
; APPLICANT: Langridge, William H. R.
; APPLICANT: Yu, Jie
; APPLICANT: Arakawa, Takeshi
; TITLE OF INVENTION: Transgenic Plant-Based Vaccines
; FILE REFERENCE: 12273-3
; CURRENT APPLICATION NUMBER: US/09/771,536B
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 15
; LENGTH: 651
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Vibrio cholerae and Escherichia coli
US-09-771-536B-15

alignment_scores:
Quality: 158.00 Length: 63
Ratio: 3.362 Gaps: 1
Percent Similarity: 74.603 Percent Identity: 50.794

alignment_block:
US-09-528-682-1 x US-09-771-536B-15 ..
Align seg 1/1 to: US-09-771-536B-15 from: 1 to: 651

188 GlyAsnSerSerArgThrIleThr..... 195
||||| ||| :|||:|||||:
460 GGAAGTATTCAGAGTAGATCTGTGAATGACTTGGATCCGTCA 509
196 ...GlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyrl 211
:|||||:|||||: |||||:|||||: |||||:|||||: |||||:|||||:
510 CATCAGTAATACCTTGGATGAGAAAAACCCAAAGTCTAGGTAAATTC 559
211 euArgGluTyrlGlnSerLysValLysArgGlnIlePheSerAspTyrlGln 227
|| |||||:|||||: |||||:|||||: |||||:|||||: |||||:|||||:
560 TTGACGAATACCAATCTTAAGTTAAAGACAAATATTTTCAGGCTATCA 609
228 SerGluValAspIleTyrlAsnArgIleArgAspGluLeu 240
|||:|||||: |||||:|||||: |||||:|||||: |||||:|||||:
610 TCTGATATGATACATACATATGATTAAGATTAAGATGAGCTTG 648

seq_name: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-771-536B-13

seq_documentation_block:
; Sequence 13, Application US/09771536B
; GENERAL INFORMATION:
; APPLICANT: Langridge, William H. R.
; APPLICANT: Yu, Jie
; APPLICANT: Arakawa, Takeshi
; TITLE OF INVENTION: Transgenic Plant-Based Vaccines
; FILE REFERENCE: 12273-3
; CURRENT APPLICATION NUMBER: US/09/771,536B
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 13


```

62 TheSerLeuSerLeuArgSerAlaHisIleAsnIleGlyGlnSerIleLeu 78
   |||::|||:::|||
370 ACTGAATCTACGTATGCCA.....GTTCGTGCAGAAAGCAATTGG 410
78 rglYrYrSeTThrYrTYrIlleYrValIleAlaThrAlaProAsnMet 95
   |||::|||:::|||
411 TGGATCTCCCTTGTTGTCATATTAAAAAT.....CTTGAACAATCT 454
95 heAsnValAsnAspValLeuGlyValYr...SerProHisProYrGlu 110
   |||::|||:::|||
455 TCACAAT.....CTTTTGCAATTCCTCCTTCATACCACACATATTAAT 498
111 GInGUValSerAlaLeuGlyGlyIleProYrSerGlnIleYrGlyTr 127
   |||::|||:::|||
499 CTTTCATCACCATTTCTTAATGGGAGTTCGGTGCACAATTG...GGAG 545
127 pTyArGrValAsnPhelGlyValIleAspGluArgLeuHisArgAsnArg 144
   |::|::|::|
546 G.....TTAATCCAATGAATTCGCACAT...CATCACT 574
144 LuTYrArGrAspArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAsp 160
   :::::|||||::|
575 TGCTCAAAAACAGATACTACTACATGATTTG....GCCAGCTATTTCGT 618
161 GlyTYrArgLeuAlaGlyPheProProAsnPhisGlnAlaTrpArgGlu 177
   |||::|||:::|||
619 GGAACTTTTTTACAAGAATC.....TCATCTGGTGTGGAAAAAGCA 662
177 uPrOTPrIleHisIleHisLa..ProGInGlyCySGlyAsnSerSerArgTr 193
   ::::|||||||::|
663 GCACATGTGTCATCACCGCAGCACAAATCTTGTGGACAGACGGAGATC 712
194 IleThrGlyAspThrCySAsnGluGluThrGlnAsnLeuSerThrIleTy 210
   ::::|::|::|::|
713 TTGATTTAGTCCCATTTCTATGCTACAGTGCAGAGACATCTCAACATTA 762
210 rLeuArGlu..... 213
   |||||
763 TCTCAGATTTCAGGGTTTANGACTTATTCAGATGGCAACATGTTCAATG 812
214 .....TYGInSerLySValLySArgGlnIlePheSer 224
   |||::|||:::|||
813 GACATTCATGCTTACCATTCCTCCCTCTCTCGTGGCTTCTTCAGT 856

seq_name: /cgn2_6/plodata/2/pna/US082_COMB.seq:US-08-245-848-1

seq_documentation_block:
: Sequence 1, Application US/08245848
: GENERAL INFORMATION:
: APPLICANT: Lochy, Camille
: APPLICANT: Lobet, Yves
: TITLE OF INVENTION: Novel Vaccine
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corp. Patents, UW2220
: STREET: P.O. Box 1539
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/245,848
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/852,212
: FILING DATE: 29-MAY-1992
: ATTORNEY/AGENT INFORMATION:

```

```

1 NAME: Sutton, Jeffrey A
2 REGISTRATION NUMBER: 34028
3 REFERENCE/DOCKET NUMBER: 12098
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: 215 270 5024
6 TELEFAX: 215 270 5090
7 INFORMATION FOR SEQ ID NO: 1:
8 SEQUENCE CHARACTERISTICS:
9 LENGTH: 4207 base pairs
10 TYPE: nucleic acid
11 STRANDEDNESS: single
12 TOPOLOGY: linear
13 MOLECULE TYPE: DNA (genomic)
14 HYPOTHETICAL: NO
15 ANTI-SENSE: NO
16 US-08-245-848-1
17
18 Alignment scores:
19 Quality: 96.50 Length: 217
20 Ratio: 1.016 Gaps: 10
21 Percent Similarity: 43.779 Percent Identity: 24.424
22
23 Alignment block:
24 US-09-528-682-1 x US-08-245-848-1 ..
25
26 Align seg 1/1 to: US-08-245-848-1 from: 1 to: 4207
27
28 5 leuTyrArgAlaAspSerArgProProAspGluIleLysArgSerGly 21
29 ::::: ||||| ||||| ||||| ::::: ::::: ||
30 627 GATACCGGTATGACTGCCGCGCCGCGGAGAGTTTCCAGAAC..GG 673
31 21 yLeuMetProArgGluHisAsnGluTyrPheAspArgGlyThrGlnMet 38
32 ::::: ||||| :::::
33 674 ATTCACGGCGTGGGAAACACAGAC..... 698
34
35 38 snLeasnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
36 ::::: ||||| ||| :::::
37 699 ....AATGTGTCGACATCTGACCGGACGTTCTCCAGTGGCGAGC 743
38 55 TyrAspAspGlyTyrAlaSerThrSerLeuSerLeuArg..... 67
39 ::::: ||||| ||||| |||||
40 744 AGCACAAGCGCTTCTGCTCCACGACGACGCGCGCGCTATACCGAGT 793
41 68 .....SerAlaHisLeuAlaGly 74
42 ::::: ||||| :::::
43 794 CTAATCGAACAATCGCATGCGAGAGCGTTGCAAGCGGCAACGCGCGCA 843
44
45 74 InsSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAlaThr 90
46 ::::: ||||| ||||| ||||| :::::
47 844 GGGGC....ACCGGCACTTCATGCGGCTACATCTACGAAGTCCGCGCC 887
48
49 91 AlaProAsnMetPheAsnValAsnAsp..ValLeuGlyValIlySerPro 106
50 |||::: ||||| :::::
51 888 GACACAATTTCTACGCGCGCGCGCAGCTGCTACTCTGCAATA..... 928
52
53 107 HisProTyrGluGlnIleValSerAlaLeuGlyGlyLeuProTyrSerGly 123
54
55 928 ..... 928
56
57 123 nileTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArgLeuH 140
58 |||
59 929 .....CGTGCAC 935
60
61 140 IsArgAsnArgGluTyrTyrArgAspArgTyrTyrArgAsnLeuAsnIleAla 156
62 |||::: ||||| :::::
63 936 ACTTATGAGGACATGCGCGCGCTATCTCGCGCGCGCTGGCGCACTTA 985
64
65 157 ProAlaGluAspGlyTyrTyrArgLeuAlaGlyPheProArgProAsnIle 173
66 ||| ::::: ||||| ||||| ||||| :::::
67 986 CCAAGAC.....CGAATATCTGGGACACGCGCG...CATTCGCC 1020
68
69 173 aTrraPdgGluGluProTrrIleHisAlaProGlnGlyCysGlyAsnS 190

```



```

1021 CCGAAAAACAT.....CCCCAGGTTAACCGCGGTC 1049
190 erserAgtThr.lleThrglYAspThrCysanSgluGluThrglInan 205
    |||||.....|||
1050 TATACACAGCGCATCCGCGGAGACCAACGACGAGATATTCAC 1097
seq_name: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-514-000-3096
seq_documentation_block:
; Sequence 3096, Application US/09514000
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)B
; CURRENT APPLICATION NUMBER: US/09/514,000
; CURRENT FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 15034
; SEQ ID NO 3096
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-514-000-3096

```

alignment_scores:

```

Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358

```

alignment_block:

```
US-09-528-682-1 x US-09-514-000-3096 ..
```

```
Align seg 1/1 to: US-09-514-000-3096 from: 1 to: 1671
```

```

1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIleY 17
  |||||.....|||
711 AACGGGCAAAACCTTCATCCGACGCTTCCAGCGCCAGCTTTCACAG 760
17 sArgSerGlyLeu...MetProArgGlyHisAsnGluThrglPheAsp 33
  : |||||.....|
761 CCTTCAGGCGGCTTTCACGCGCGCGGCGGCGAAT.....A 798
33 rg.glyThrglMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrgl 49
  |||||.....||
799 GCGCGTGTATGCGCGCATATCGGCATC.....CGCAGCTTCA 836
49 nThrglYpHeValArgTYrAspAspGlyTYrValSerThrSerLeuSerL 66
  | ||| :
837 ATTGCGGTGTTTTCAG.....CTCCGCG 859
66 euArgSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTYr..... 80
860 AATTCACACCTGACCTGCGGCGCAGTTCCTTCGCGCGGTTTCATGAC 909
81 .....SerThrglTYrIleTYrValIleAlaThrglAlaProAsnMe 94
  |||||.....|||
910 GAGATCTGTATGCGCGCTGTATGTTGTGTGTCGCGCGCGCGCAGGT 959
94 tPheAsnValAsnAspValLeuGlyValTYrSerProHisProTYrGluG 111
  : |||||.....|
960 TATTAACGCGCATGACGCGCTGTATAGACACAGGTCCTCTCCGCG 1009
111 lngluValSerAlaLeuGlyGly.....IleProTYr 121
  |||||.....|||
1010 AGGAGTGCAGAAATCTTGTGGCATGATGCGCGTACGCCCTGCCGCC 1059
122 SerGlnIleTYrGly.....TrpTYrArgVal 130
  : |||
1060 GCGCAGCTAAACGCAAAACGATCTCACCCCTTGTCTTCTTTTCAGAT 1109
130 lAsnPhcGlyValIleAspGluArgLeuHisArgAsnArgGluTYrArg 147

```

```

1110 GCAGGCGCATATCTGTGAC.....CTGAATGGCGCGCGTCCGCTCG 1153
147 sPrArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAspGlyTYrArg 163
  |||||.....|||
1154 GCACCCGTCACCCGCGATCTTGAACTTCGCGCGGACAGACAGAAATCCG 1203
164 leuAlaGlyPheProProAspHisGln.....AlaTrpArgGluG 177
  : ||| |||
1204 ATGACGCAAGAACCATTTGACGCGAGATTTCCGATAGGACGCGGATCGG 1253
177 wProTPrIleHisAlaProGlnGlyCysGlyAsnSerSerArgThrl 194
  : |||||.....|
1254 CGACTTCGTC.....TCGCGCACACCCCGGAATG 1285
194 lerThrglYAspThrCysAsnGluGluThrglAsnLeuSerThrlIleTYr 210
  || |||
1286 TCGCGCGTCACTTCGCAATGCAGTTACCGGACGAGTCTGCAGTCCGTCAT 1335
211 leuArg.....GluTYrGlnSerLysValLysArg 220
  : |||||.....|
1336 TTCACGCTTGCCAGTTCCAGAGAGATATCCGCGT 1371

```

```
seq_name: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-739-449-5019
```

seq_documentation_block:

```
; Sequence 5019, Application US/09739449
```

GENERAL INFORMATION:

```

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 5019
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-5019

```

alignment_scores:

```

Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358

```

alignment_block:

```
US-09-528-682-1 x US-09-739-449-5019/rev ..
```

```
Align seg 1/1 to reverse of: US-09-739-449-5019 from: 1 to: 1671
```

```

1 AsnGlyAspArgLeuTYrArgAlaAspSerArgProProAspGluIleY 17
  |||||.....|||
961 AACGGGCAAAACCTTCATCCGACGCTTCCAGCGCGGCTTTCACAG 912
17 sArgSerGlyLeu...MetProArgGlyHisAsnGluThrglPheAsp 33
  : |||||.....|
911 CCTTCAGGCGGCTTTCACGCGCGCGGCGGCGAAT.....A 874
33 rg.glyThrglMetAsnIleAsnLeuTYrAspHisAlaArgGlyThrgl 49
  |||||.....||
873 GCGCGTGTATGCGCGCATATCGGCATC.....CGCAGCTTCA 836
49 nThrglYpHeValArgTYrAspAspGlyTYrValSerThrSerLeuSerL 66
  | ||| :
835 ATTGCGGTGTTTTCAG.....CTCCGCG 813
66 euArgSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTYr..... 80
  : |||||.....|
812 AATTCACCTGACCTGCGCGGCGCAGTTCCTTCGCGCGGTTTCATGAC 763

```


[illegible]

```
seq_documentation_block:
  Sequence 1915, Application US/60168139
  GENERAL INFORMATION:
    APPLICANT: Hinkle, Gregory J.
    APPLICANT: Slater, Steven C.
    TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequence And Uses Thereof
    FILE REFERENCE: 38-21(15490)A
    CURRENT APPLICATION NUMBER: US/60/168,139
```



```

: CURRENT FILING DATE: 1999-12-01
: NUMBER OF SEQ ID NOS: 3432
: SEQ ID NO 1915
: LENGTH: 4199
: TYPE: DNA
: ORGANISM: Agrobacterium tumefaciens
US-60-168-139-1915

```

Quality:	96.00	Length:	246
Ratio:	0.768	Gaps:	12
Percent Similarity:	50.813	Percent Identity:	22.358

alignment_block:

US-09-528-682-1 x US-60-168-139-1915/rev .

Align seg 1/1 to reverse of: US-60-168-139-1915 from: 1 to: 4199

[illegible]

211 LeuArg.....GluTrpInSerLysValLysArg 220
 ::::: :
 853 TTCACGCTTGCACGTTTCAGAGAGGATATCCGGCCT 818

seq_name: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-620-392-45559

seq_documentation_block:

; Sequence 45559, Application US/096620392

; GENERAL INFORMATION:

APPLICANT: Boukharov, Andrey A
ADDITIONAL: David K

APPLICANT: KOVALLIC, David K.

APPLICANT: Liu, Jinguang

FILED: TITLE OF INVENTION: Plate

FILE REFERENCE: 38-21(5)

CURRENT APPLICATION NUMBER: U

; CURRENT FILING DATE: 2000-07-19

; NUMBER OF SEQ ID NOS: 6965.

; SEQ ID NO 45559
TENCSTU 7413

LENGTH: 7412
TYPE: DNA

ORGANISM:

US-09-620-392

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

US-09-620-392-45559

alignment_scores:

Quality:	96.00	Length:	195
----------	-------	---------	-----

Ratio:	1.103	Gaps:	3
Percent similarity:	44.615	Percent identity:	24.103

Percent similarity: 44.013 Percent identity: 24.103

alignment_block:
PS-09-528-682-1

00004-75C-070-60-SO Y T 790-87C-60-SO

Align seg 1/1 to: 05-09-020-392-45539 L10011: 1 to: /412

```

2  G|VAsPArGLuE|T|YrAlA|AsPSeR|ArPrRoAsP|G|U|L|E|Yr|SAr 18
21  ||||| ||||| ||||| ||||| |||||
25  GGTGTCAAGGCTTACGGACACAGATTCCAG.....AAGAG 27/97
18  gserGlycylLeuMetProArgGlyHisAsnGlyTyrPhe..... 31
1  ||||| ||||| ||||| |||||
280 AAGAGAGAGGTTGGGTGGCGGGCGGGCCAGTGAATPACTTTAGTGGCG 328/280
32  ....AspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArg 46
1  ||||| ||||| ||||| |||||
330 ACCCAGATGTGGCGCTCATGCAGACACCTGGCGGACGGACATGCTCGG 379/330
47  G|Yr|Hr|G|InT|Hr|G|Yr|Phe|Val|ArG|Tyr|Asp.....AspG|Yr|Val|Se 61
11  ||||| ||||| ||||| |||||
380 GGGCACCCTGATGGCTGGCGGGCCACCCAGTCATCTTGAAGGTAACCTGC 429/380
61  rThrSerLeuSerLeuArGSerAlaHisLeuAlaG|Yr|G|InSer|Leu|S 78
1  : ||||| ||||| ||||| |||||
430 CCTGGCCCTGGCGCTGGCGGGCGGCGACAGACTGGCCAG..... 465/430
78  eRg|Yr|Ser|Thr|Tyr|T|Leu|Yr|Val|Ile|Ala|Hr|Ala|Pro|Asn|Met 94
1  : ||||| ||||| ||||| |||||
465  ..... 465/465
95  PheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGluG| 111
1  : ||||| ||||| ||||| |||||
470  .....CTTCCACATGGCGTGTCTCGA 489/470
111  nG|U|Val|Ser|Ala|Leu|G|Yr|Ile|Pro|Tyr|Ser|G|In|Leu|Tyr|Prt 128
1  : ||||| ||||| ||||| |||||
490 GCGCCTCCACACAGGTCGCGGCGCTC.....GAGGTCTGTGGC.... 526/490
128  YrArG|Val|Asn|Phe|Val|Val|Ile|AsP|Yr|G|U|L|E|Yr|SAr|G|U| 144
1  ||||| ||||| ||||| |||||
527  .....GATCTCGAGGTGCGCGGCGCGCGGCGG 553/527

```



```
145 TyrArgAsp...ArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAs 160
:::|||||
554 CACCGAGACCCCGCGGTATGCGCGCGGAGACGACGTGCGCCCGCTCCACC 603
|||||
160 pGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTrpArgGluG 177
|||||
604 GAGTGTGTCTGTGTCGCGCTTCCCGCACACGACGCC.....GAGC 647
177 IuProTrpIleHisHis.....AlaProGln 185
|||
648 CGCCACCCCGCCCATCGACGGTCCCGCCGACACGCC 682

seq_name: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-620-392-59272

seq_documentation_block:
; Sequence 59272, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIntosh, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 59272
; LENGTH: 7657
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-59272
```

```
alignment_scores:
Quality: 96.00 Length: 195
Ratio: 1.103 Gaps: 9
Percent Similarity: 44.615 Percent Identity: 24.103
```

```
alignment_block:
US-09-528-682-1 x US-09-620-392-59272 ..
Align seg 1/1 to: US-09-620-392-59272 from: 1 to: 7657
```

```
2 GtAspArgLeuTyrArgAlaAspSerArgProProAspGluIleuysat 18
|||||
5145 GGCTGACGCTACGAGCAAGATTCACAG.....AAGAG 5179
18 gSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPhe..... 31
|||||
5180 AAGAAAGAGGTTGGTGGCGCGCGCGCATGATACCTTGATGAGTGGCG 5229
32 ....AspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArg 46
|||||
5230 ACCCAGATGTCGCGCTGACGACACCTGCGCGGACGACATGCTCGG 5279
47 GtThrGlnThrGlyPheValArgTyrAsp.....AspGlyTyrValSe 61
|||||
5280 GGGCACCCTGATGCGCTGGCGGACGCGAGATCTTGACGTAAGTCCG 5329
61 rTrpSerLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleuS 78
|||||
5330 CCGCGCGCTGCGCGTGGCGCGCGCGACGACGCTGGCGCAG..... 5369
78 erGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMet 94
5369 .....
95 PheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGluG 111
5370 .....CTTCACATGCGGTCGTGCA 5389
111 nGluValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTrp 128
.....
```

```
5390 GCGCTCTCCACAGGTGCGCGCGCTC.....GAGTGTGAGGG..... 5426
128 yTrpValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGlu 144
|||||
5427 .....GATCTCGAGGTGCGCGCGCGCGCGCGCGCG 5453
145 TyrArgAsp...ArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAs 160
:::|||||
5454 CACCGAGACCCCGCGGTATGCGCGCGGAGACGACGTGCGCCCGCTCCACC 5503
160 pGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTrpArgGluG 177
|||||
5504 GAGTGTGTCTGTGTCGCGCTTCCCGCACACGACGCC.....GAGC 5547
177 IuProTrpIleHisHis.....AlaProGln 185
|||
5548 CGCCACCCCGCCCATCGACGGTCCCGCCGACACGCC 5582

seq_name: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-514-000-219
```

```
seq_documentation_block:
```

```
; Sequence 219, Application US/09514000
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Hunkle, Gregory J.
```

```
; APPLICANT: Slater, Steven C.
```

```
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
```

```
; FILE REFERENCE: 38-10(15490)B
```

```
; CURRENT APPLICATION NUMBER: US/09/514,000
```

```
; CURRENT FILING DATE: 2000-02-23
```

```
; NUMBER OF SEQ ID NOS: 15034
```

```
; SEQ ID NO 219
```

```
; LENGTH: 11719
```

```
; TYPE: DNA
```

```
; ORGANISM: Agrobacterium tumefaciens
```

```
US-09-514-000-219
```

```
alignment_scores:
Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358
```

```
alignment_block:
US-09-528-682-1 x US-09-514-000-219 ..
```

```
Align seg 1/1 to: US-09-514-000-219 from: 1 to: 11719
```

```
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleu 17
|||||
6103 AACGGGCAAAACCTTCATCCGCAACGCTTCCAGCGCGCATGCTTCCAG 6152
17 sArgSerGlyGlyLeu...MetProArgGlyHisAsnGluTyrPheAsp 33
:::|||||
6153 CTTTCAGCGCGGTCTTCAACGCGCGCGCGCGCGCAAT.....A 6190
33 rGtGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrG 49
|||||
6191 GCGCGCTATGCGCGCAATATCGGCATC.....CGCAGCTTCA 6228
49 nTrpGlyPheValArgTyrAspArgGlyTyrValSerThrSerLeuSer 66
|||||
6229 ATTGCGGTGTTTTCAG.....CTCCGCG 6251
66 euArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyr..... 80
:::|||||
6252 AATTCAACCTGCACTGCGCGCGCGCAGTCTTTCGCGGTTTCATGATAC 6301
81 .....SerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMe 94
|||||
6302 GAGATCTTCATGCGCGCTTGTATTTGTTGTGCGCGCGCGCTGCAAGT 6351
94 tPheAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGluG 111
.....
```


6352 TATAAGCGCGCATGACGCCGTGTTAGACACAGAGTCTCTCCGCC 6401
111 InGUValSerAlaLeuGly.....IleProTyr 121
|||||.....
6402 AGGAACTCGCGAATCTTCTGGCATCATGGCGTACGCCCTGCCGCC 6451
122 SerGlnIleTyrGly.....TrpTyrArgVa 130
.....
6452 GCCGACCTAAACGCAAAACCGATCTCACCCCTGCTCTTTCAGAT 6501
130 LasnPheGlyValIleAspGluArgLeuHisArgAsnArgGluTyrArg 147
.....
6502 GCAGCGCGATATCGTGAC.....CTGAATGGGGGGCGCGTCCGCTCG 6545
147 sPARgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArg 163
.....
6546 GCACCCGTCACCGGATCTTGAACTGCGGCGCAGAACGAAATTCGG 6595
164 LeuAlaGlyPheProAspHisGln.....AlaTPrArgGluG 177
.....
6596 ATGCACGGAAGACCATGACGCGATTCGCATAGGACGCGGATCGG 6645
177 uProTPrIleHisIleAlaProGlnGlyCysGlyAsnSerSerArgThr 194
.....
6646 CGACTTCGTC.....TGCAGCAGCACCCGCAAAATGA 6677
194 leThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyr 210
.....
6678 TCGGCGGTCACGTTGCGAATGCAATTACCGACGCTCTGCAAGTGGTGCAT 6727
211 LeuArg.....GluTyrGlnSerLysValLysArg 220
.....
6728 TTCACGCTTGCCAGTTCAGAGAGATATCCGGCGT 6763

[illegible]

/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-105-229-7302	-	71.00	111.93	403.7	
/cgn2_6/p.todata1/pna/PCT_NEW_COMB.seq:PCT-US02-09288-315	-	71.00	111.18	443.24	
/cgn2_6/p.todata1/pna/PCT_NEW_COMB.seq:PCT-US02-09288-35	-	71.00	110.52	482.56	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-10-667-1170A-422	+	71.00	100.16	1.6e+07	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-105-299-12480	-	70.50	128.10	50.5	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-027-63-216230	+	70.50	121.49	118.1	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-54-020B-9250	+	70.50	120.91	127.7	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-360-03-14352	+	70.50	120.84	128.1	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-91-002-12498	+	70.50	118.20	180.5	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-60-360-03-40452	+	70.50	117.96	185.0	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-40734	+	70.50	116.70	218.7	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-23551	+	70.50	116.70	218.7	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-104-047-1207	+	70.50	110.72	479.53	
/cgn2_6/p.todata1/pna/PCT_NEW_COMB.seq:PCT-US02-13142-568	-	70.50	110.72	470.53	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-128-714-5568	-	70.50	110.68	472.5	
/cgn2_6/p.todata1/pna/PCT_NEW_COMB.seq:PCT-US02-13142-5568	-	70.50	110.68	472.5	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-128-714-5568	-	70.50	108.08	659.91	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-088-697-1	+	70.50	104.36	1.1e+03	
/cgn2_6/p.todata1/pna/US08_NEW_COMB.seq:US-08-793-279C-1	+	70.50	104.22	1.1e+03	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-919-02-9441	+	70.50	103.99	1.1e+03	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-05-372B-1354	+	70.50	103.99	1.1e+03	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-100-912-3	+	70.50	103.99	1.1e+03	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-09-935-625-17845	+	70.50	101.04	1.6e+06	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-935-625-28236	+	70.50	101.04	1.6e+06	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-150-407-1	+	70.50	96.10	3.1e+03	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-027-63-76212	-	70.50	97.42	1.1e+06	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-027-63-242617	-	70.00	128.07	50.1	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-60-360-03-93951	+	70.00	127.14	57.1	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-33951	+	70.00	123.62	89.8	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-33951	+	70.00	123.62	89.8	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-152-886-52	+	70.00	123.39	92.6	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-46955	+	70.00	116.83	214.3	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-104-04-1125	+	70.00	114.92	301.5	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-982-667-2	+	70.00	114.02	308.06	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-09-94-024-24-2	+	70.00	114.02	308.06	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-10-033-247-2	+	70.00	114.02	308.06	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-033-247-2	+	70.00	112.60	369.1	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-09-54-020B-9687	+	70.00	112.60	369.1	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-105-229-12682	-	70.00	99.16	2.1e+06	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-939-968A-1	-	70.00	60.41	2.8e+05	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-027-63-277332	+	69.50	126.89	59.1	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-027-63-207319	+	69.50	126.83	59.1	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-34268	+	69.50	125.32	72.3	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-34709	+	69.50	121.66	115.1	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-37537	+	69.50	121.05	125.1	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-26055	+	69.50	120.53	133.3	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-46259	+	69.50	120.15	140.3	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-089-64-13	+	69.50	118.55	17.84	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-424898	+	69.50	118.05	183.1	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-382573	+	69.50	117.74	191.1	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-31983	+	69.50	117.53	196.1	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-239370	+	69.50	112.77	361.1	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-239370	+	69.50	112.77	361.1	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-105-229-3066	+	69.50	111.90	404.4	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-074-04-65	+	69.50	104.74	1.0e+03	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-105-299-12469	-	69.50	98.59	2.2e+06	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-70-360-03-41968	-	69.50	113.65	323.17	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-24384	+	69.50	113.58	326.1	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-994-404-155	-	69.50	97.76	2.8e+03	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-994-404-155	-	69.50	96.74	2.8e+03	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-994-404-2628	-	69.50	131.07	4.8e+03	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-919-002-1617	-	69.50	126.70	60.62	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-602-777A-257	+	69.00	123.46	91.4	
/cgn2_6/p.todata1/pna/US09_NEW_COMB.seq:US-09-602-777A-257	+	69.00	123.46	91.4	
/cgn2_6/p.todata1/pna/US10_NEW_COMB.seq:US-10-027-63-23170	-	69.00	123.69	89.5	
/cgn2_6/p.todata1/pna/US60_NEW_COMB.seq:US-60-360-03-42764	+	69.00	123.49	89.5	

[illegible]

[illegible]

cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-13142-560 +	66.00	103.36	1.2e+03	3
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-540-2099-809 +	66.00	103.26	1.2e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-112952 +	66.00	103.21	1.2e+03	3
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-13142-5535 +	66.00	103.20	1.2e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-128-714-5535 +	66.00	103.12	1.2e+03	3
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-053-375B-4470 +	66.00	102.85	1.3e+03	3
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-442-884B-468 +	66.00	102.85	1.3e+03	3
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-13142-151 +	66.00	102.85	1.3e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-128-714-151 +	66.00	102.83	1.3e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-113891 -	66.00	102.70	1.3e+03	3
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-13142-449 +	66.00	102.63	1.3e+03	3
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-13142-5151 +	66.00	102.63	1.3e+03	3
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-13142-5449 +	66.00	102.63	1.3e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-128-714-449 +	66.00	102.63	1.3e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-128-714-451 +	66.00	102.63	1.3e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-128-714-5449 +	66.00	102.63	1.3e+03	3
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-13142-535 +	66.00	102.36	1.4e+03	3
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-27774 +	66.00	102.22	1.4e+03	3
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-13142-5560 +	66.00	102.13	1.4e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-128-714-5560 +	66.00	102.13	1.4e+03	3
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-116-802-85 +	66.00	98.81	2.2e+03	49
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-105-299-12472 -	66.00	96.89	2.8e+03	43
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-105-299-12545 -	66.00	96.88	2.8e+03	43
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-105-299-12568 -	66.00	96.85	2.8e+03	43
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-09357-8619 +	66.00	96.06	3.1e+03	6
cgml2_6/p/odataa/1/pna/PCT_NEW_COMB.seq	:PCT-US02-09357-8711 +	66.00	96.06	3.1e+03	6
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-105-299-12458 +	66.00	96.06	3.1e+03	6
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-104-045-72 -	66.00	86.98	9.9e+03	10
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-107-331-280 -	66.00	76.18	3.9e+04	4
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-939-964A-1 +	66.00	52.13	7.0e+05	53
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-161613 -	65.50	125.60	69.79	80
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-919-912-021-30 +	65.50	115.98	239.60	80
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-365-264-527 +	65.50	112.74	100.66	45
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:PCT-US01-43704-32 +	65.50	121.79	113.73	45
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-721-544-5286 +	65.50	120.47	113.80	45
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-128276 +	65.50	118.84	166.02	4
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-128276 +	65.50	118.52	173.07	4
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-128276 +	65.50	118.16	181.26	4
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-223618 -	65.50	118.16	181.26	4
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-223618 -	65.50	116.91	212.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-30949 -	65.50	116.91	212.57	4
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-360-039-28191 -	65.50	116.36	228.25	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-09-712-021-25 +	65.50	115.98	239.60	80
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-919-912-021-8182 +	65.50	115.98	239.60	80
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-835-604-907 +	65.50	113.48	292.08	80
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-835-604-909 +	65.50	113.18	343.02	1
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-39225 -	65.50	112.25	386.79	1
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-39256 -	65.50	112.25	386.79	1
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-39967 -	65.50	112.25	386.79	1
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-39987 -	65.50	112.25	386.79	1
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-540-2098-2602 +	65.50	112.12	393.14	4
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-540-2098-2602 +	65.50	111.92	403.34	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-34540 +	65.50	111.84	407.60	4
cgml2_6/p/odataa/1/pna/US10_NEW_COMB.seq	:US-10-027-632-240031 -	65.50	111.16	444.70	4
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-935-625-20976 +	65.50	111.16	444.70	4
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-935-625-20976 +	65.50	111.16	444.70	4
cgml2_6/p/odataa/1/pna/US09_NEW_COMB.seq	:US-09-935-625-20976 +	65.50	111.16	444.70	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	111.11	447.31	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	111.07	461.75	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4
cgml2_6/p/odataa/1/pna/US60_NEW_COMB.seq	:US-60-360-039-428935 +	65.50	109.59	543.57	4

[illegible]

/cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-128-714-6307 + 64.50 116.54 222.93
/cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-149787 - 64.50 114.79 279.14
/cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-33329 + 64.50 112.99 351.63
/cgn2.6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-33742 - 64.50 112.37 380.89
/cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-704-302A-280 + 64.50 112.35 381.92
/cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-3716 + 64.50 112.24 387.11
/cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-11904 - 64.50 112.19 389.71
/cgn2.6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-41713 + 64.50 112.08 394.93
/cgn2.6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-43416 + 64.50 112.05 396.49
/cgn2.6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-37252 + 64.50 110.57 479.75
/cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-122562 + 64.50 109.23 569.70

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-809-033A-1

seq_documentation_block:

; Sequence 1, Application US/09809033A

; GENERAL INFORMATION:

; APPLICANT: McGhee, Jerry

; Kiyono, Hiroshi

; Takeda, Yoshitumi

; Ohmura, Mari

; Yamamoto, Shingo

; TITLE OF INVENTION: Chimeric Nontoxic Mutants of

; Enterotoxins as Mucosal Adjuvants for Cell-Mediated or

; Humoral Immunity

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hendricks and Assoc

; STREET: P.O. Box 2509

; CITY: Fairfax

; STATE: VA

; COUNTRY: US

; ZIP: 22031

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/809,033A

; FILING DATE: 16-Mar-2001

; CLASSIFICATION: <unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Hendricks, Glena

; REGISTRATION NUMBER: 32,535

; REFERENCE/DOCKET NUMBER: MCG-01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703/425-8405

; TELEFAX: 703/425-8406

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2022 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-809-033A-1

alignment_scores:

Quality: 1053.00 Length: 240

Ratio: 4.701 Gaps: 0

Percent Similarity: 93.33 Percent Identity: 80.833

alignment_block:

US-09-528-682-1 x US-09-809-033A-1

Align seg 1/1 to: US-09-809-033A-1 from: 1 to: 2022

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspLulely 17

|||||
572 AATGATCATTAATTAATCGGACGATTTCTAGACCTTCATCAATAATAA 621
17 sArgSerGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
622 GCAGTCAGGTGGCTTATGCCAAGACCACAGAGTGAATCTTGACCGAG 671
34 1YhrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
672 GTACTCAATGATATCAACCTTATGATCATGCACAGAGAACTCAGACG 721
51 GlyPheValArgTyrAspGlyTyrValSerThrSerLeuSerLeuArg 67
722 GGATTTGTTAGGCACGATGATGATGTTCCACCTCAATAGTTGTAG 771
67 gSerAlaHisLeuAlaGlyInsSerIleLeuSerGlyTyrSerThrTyr 84
772 AAGTCCCTTGTAGTGGTCAACTATATTGTGTGTCATTTACTTATT 821
84 yrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
822 ATATATATATGTTATAGCCACGACCCACACATGTTTAACGTTATATGA 870
101 LeuGlyValTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
871 TTAGGGGCATACAGCTCCTCATCCAGATGAAAGAGTTCTCTTAGG 920
117 yGlyIleProTyrSerGlnIleTyrGlyTyrPyrArgValAspPheGly 134
921 TGGGATTCATCTCCCAATATATGATGATGATTCGATTCATTTTGGG 970
134 allleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyr 150
971 TGCTGTATGAACAATTATACATCGTAATAGGGGCTACAGAGATATATTC 1020
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGly 167
1021 AGTAACCTTATATGCTCCACAGCAGAGATGATGATGATGATGATGAT 1070
167 eProProAspHisGlnAlaThrPargGluGluProTyrPheHisAlaP 184
1071 CCTCCGACATACAGCTTGGAGGACAGACCGCTGATTCATCATACGAC 1120
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
1121 CGCGGGTGTGGGATGCTCAAGATCATCATGATGATGATGATGATGAT 1170
201 GlnGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
1171 GAAAAAACCCAAAGCTAGGTGTAATAATCTTGACGAAATCAACAACTCA 1220
217 sValIleAspGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
1221 AGTTAAAGACAAATATTTTCAGAGCTATCATCTGATATTGATACACAT 1270
234 snArgIleArgAspGluLeu 240
1271 ATGAATTAAGATGATTA 1290

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-809-033A-2

seq_documentation_block:

; Sequence 2, Application US/09809033A

; GENERAL INFORMATION:

; APPLICANT: McGhee, Jerry

; Kiyono, Hiroshi

; Takeda, Yoshitumi

; Ohmura, Mari

; Yamamoto, Shingo

; TITLE OF INVENTION: Chimeric Nontoxic Mutants of

; Enterotoxins as Mucosal Adjuvants for Cell-Mediated or

; Humoral Immunity

; NUMBER OF SEQUENCES: 4


```

33 rg.glythrglmetasnilleasnleuylrpsphisaalargglythrgl 49
   |||
   |||||
873 GCGGTGATGCGCGCAATATCGCATC.....CGGCAAGCTTCA 836
   |||
49 ntrhglyphevalargtyrpsaspglytyrvalserthrserserleu 66
   |||
835 ATTCGGGTGTTTCAG.....CTCGCG 813
66 euargseralahnisleuaglinsertleuaserglytyr..... 80
   |||||
812 AATTCACCTCGACCTGGCGGCGAGTTCTTCTGCGCGGTTATGATGAC 763
81 .....Serthrytyriletyrvallealathralaproasme 94
   |||||
762 GAGATCTGTATGCGCGCTGTATTTGTTGTCGCGGCGCGTCAAGT 713
94 tpheasnvalasnapsvalleuglyvaltyrserprohisprotyr 111
   |||||
712 TATTAACGCGCATGACGCGCGGTATAGACAGCAGTCTCTTCGCGC 663
111 lngluvalseralaleuglygly.....lleprotyr 121
   |||||
662 AGGAGTGGCGAATCTTCTTGCGATATGCGCTACGCGCTCCGCC 613
122 serglnletyrgly.....trytyrargva 130
   |||||
612 GCCGACGTAAACGGCAAAACCGATCTCACCTTGTCTTTTTCAGAT 563
130 lasnphneglyvalilleaspgluargleuhsargasnargglytyr 147
   |||||
562 GCAGGCGCATATCGTGAC.....CTGAATGGCGCGCGCTCCGCTCG 519
147 spargtyrtyrargasnleuasnillealaproalagluasp 163
   |||||
518 GCACCCGTCACCGGATCTTAATCTGCGCGCGGCAAGCAATTCGCG 469
164 leuaglapherpropoasphisgln.....alatprargglugl 177
   |||||
468 ATGCACGAGAACCATGACGAGATTTCCGATAGGACGCGGATCGG 419
177 uPrtprlehislsalaproglnglycysglyasnserserarghrt 194
   |||||
418 CGACTTGTC.....TCGGCAGCACCGCGCAATGA 387
194 lethrglyaspthrcysangluu thr gl n as n leu ser th r l e ty r 210
   |||||
386 TCGGCGGTGACGTTCGCAATGCACTTACCGCAGCTCTGCACTGCGT 337
211 leuarg.....glu ty r g l n s e r l y s v a l l y s a r g 220
   |||||
336 TTCACGCTTGCCAGTTCAGAGAGATATCCGGCGT 301
seq_name: /cgn2_6/prodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-38194
seq_documentation_block:
; Sequence 38194, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360, 039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38194
; LENGTH: 1653
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens

```

```

US-60-360-039-38194
alignment_scores:
  Quality: 96.00      Length: 246
  Ratio: 0.768      Gaps: 12
  Percent Similarity: 50.813      Percent Identity: 22.358
alignment_block:
US-09-528-682-1 x US-60-360-039-38194/rev ..
Align seg 1/1 to reverse of: US-60-360-039-38194 from: 1 to: 1653

1 asnglyspargleuylrargalaaasparargproprospgluleuyl 17
   |||||
961 AACGCGAAACCTTCAATCCGACGCTCCAGCGCGGCTTCCGAG 912
17 sarqserglyglyleu...metproargglyhisanglu ty r p h e a s p a 33
   |||||
911 CATTACGCGCGCTTTCACGCGCGCGCGCGGCAAT.....A 874
33 rg.glythrglmetasnilleasnleuylrpsphisaalargglythrgl 49
   |||||
873 GCGGTGATGCGCGCAATATCGCATC.....CGGCAAGCTTCA 836
49 ntrhglyphevalargtyrpsaspglytyrvalserthrserserleu 66
   |||||
835 ATTCGGGTGTTTCAG.....CTCGCG 813
66 euargseralahnisleuaglinsertleuaserglytyr..... 80
   |||||
812 AATTCACCTCGACCTGGCGGCGAGTTCTTCTGCGCGGTTATGATGAC 763
81 .....Serthrytyriletyrvallealathralaproasme 94
   |||||
762 GAGATCTGTATGCGCGCTGTATTTGTTGTCGCGGCGCGTCAAGT 713
712 TATTAACGCGCATGACGCGCGGTATAGACAGCAGTCTCTTCGCGC 663
111 lngluvalseralaleuglygly.....lleprotyr 121
   |||||
662 AGGAGTGGCGAATCTTCTTGCGATATGCGCTACGCGCTCCGCC 613
122 serglnletyrgly.....trytyrargva 130
   |||||
612 GCCGACGTAAACGGCAAAACCGATCTCACCTTGTCTTTTTCAGAT 563
130 lasnphneglyvalilleaspgluargleuhsargasnargglytyr 147
   |||||
562 GCAGGCGCATATCGTGAC.....CTGAATGGCGCGCGCTCCGCTCG 519
147 spargtyrtyrargasnleuasnillealaproalagluasp 163
   |||||
518 GCACCCGTCACCGGATCTTAATCTGCGCGCGGCAAGCAATTCGCG 469
164 leuaglapherpropoasphisgln.....alatprargglugl 177
   |||||
468 ATGCACGAGAACCATGACGAGATTTCCGATAGGACGCGGATCGG 419
418 CGACTTGTC.....TCGGCAGCACCGCGCAATGA 387
177 uPrtprlehislsalaproglnglycysglyasnserserarghrt 194
   |||||
194 lethrglyaspthrcysangluu thr gl n as n leu ser th r l e ty r 210
   |||||
386 TCGGCGGTGACGTTCGCAATGCACTTACCGCAGCTCTGCACTGCGT 337
211 leuarg.....glu ty r g l n s e r l y s v a l l y s a r g 220
   |||||
336 TTCACGCTTGCCAGTTCAGAGAGATATCCGGCGT 301

```



```

seq_name: /cgn2_6/plodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-38598
seq_documentation_block:
: Sequence 38598, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: FILE REFERENCE: 38-10(52052)A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 38598
: LENGTH: 1653
: TYPE: DNA
: ORGANISM: Agrobacterium tumefaciens
US-60-360-039-38598

```

```

alignment_scores:
Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358

```

alignment_block:

US-09-528-682-1 x US-60-360-039-38598/rev ..

Align seg 1/1 to reverse of: US-60-360-039-38598 from: 1 to: 1653

```

1 ASnglYAspArgLeuTYrArgAlaAspSerArgProProAspIuileY 17
: |||||:|||||:|
961 AACGGGCAAAACCTTATCCGACGCTCCAGCGCGGCTTCCAG 912
17 sArgSerGlyLeu...MetProArgLynHisAsnGluTYrPheAsp 33
: :|||||:|
911 CTTTCAGGGGGCTTTTCAGCGCGGGCGGCGAAAT.....A 874
33 rg.gLYThrcImetAsnIleAsnLeuTYrAspHisAlaArgLYThrg 49
: |||||:|||||:|
873 GGGGTGATGGCGGCAATATCGGCATC.....CGGACGTTCA 836
49 nThrgLYPheValArgTYrAspAspIuTYrValSerThSerLeuSer 66
: |||||:|
835 ATCGGGTGTTCAG.....CTCGCG 813
66 euArgSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTYr..... 80
812 AATTCAACSTCGACSTGGGGCGGCACTTCTGCGCGGTTTCANGTAC 763
81 .....SerThTYrTYrIleTYrValIleAlaThAlaProAsnme 94
: |||||:|||||:|
762 GAAATCTTGATCGCGCCTTGATTTGTGTGCGCGCGCGCGT 713
94 tPheAsnValAsnAspValLeuGlyValTYrSerProHisProTYrGlu 111
: |||||:|||||:|
712 TATAAACGGCGATGACGCGGTGTATAGACAGCAGGTCTCTCGCG 663
111 lngIValSerAlaLeuGlyLY.....IleProTYr 121
: |||||:|
662 AGGAATCGGGAATCTTGGGATCATCGGCTACGCGCGCGCGCC 613
122 SerGlnIleTYrGly.....TrpTYrArgVA 130
: |||||:|
612 GCGGAGTAAACGGCAAAACCGATTCACCTTGCTTTTTCGAT 563
130 lAsnPheGlyValIleAspGlyArgLeuHisArgAsnArgGlyTYrArg 147
: |||||:|
562 GCAGGGCGGATATCGTGAC.....CTGAATGCGGGCGGTGCGCGTGC 519

```

```

147 sPArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAspIuTYrArg 163
: |||||:|||||:|
518 GCACCCGTCACCCGATCTTGAATCTCGCGGCGAGACAGCAATTCGG 469
164 LeuAlaGlyPheProProAspHisGln.....AlaTrpArgGlu 177
: |||||:|
468 ATCAGCGAAGACCATTTGACGAGATTTCCGATAGGAGCGGATTCGG 419
177 uProTrpIleHisAlaProGlnGlyCysGlyAsnSerSerArgThr 194
: |||||:|
418 CGACTTCGTC.....TCCGGCAGCACCGCGGAATGA 387
194 lPheThGlyAspThrcYAsnGluGluThrcGlnAsnLeuSerThrIleTYr 210
: |||||:|
386 TCGCGGCTCACGTTGCGAATGCAATTACCGACGCTGCGATCGCGAT 337
211 LeuArg.....GluTYrGlnSerLYsValysArg 220
: |||||:|
336 TTCACGCTTGCACAGTTTCAGAGAGATATCCGGCGT 301

```

seq_name: /cgn2_6/plodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-37962

```

seq_documentation_block:
: Sequence 37962, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: FILE REFERENCE: 38-10(52052)A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 37962
: LENGTH: 1657
: TYPE: DNA
: ORGANISM: Agrobacterium tumefaciens
US-60-360-039-37962

```

```

alignment_scores:
Quality: 96.00 Length: 246
Ratio: 0.768 Gaps: 12
Percent Similarity: 50.813 Percent Identity: 22.358

```

alignment_block:

US-09-528-682-1 x US-60-360-039-37962/rev ..

Align seg 1/1 to reverse of: US-60-360-039-37962 from: 1 to: 1657

```

1 ASnglYAspArgLeuTYrArgAlaAspSerArgProProAspIuileY 17
: |||||:|||||:|
961 AACGGGCAAAACCTTATCCGACGCTCCAGCGCGGCTTCCAG 912
17 sArgSerGlyLeu...MetProArgLynHisAsnGluTYrPheAsp 33
: :|||||:|
911 CTTTCAGGGGGCTTTTCAGCGCGGGCGGCGAAAT.....A 874
33 rg.gLYThrcImetAsnIleAsnLeuTYrAspHisAlaArgLYThrg 49
: |||||:|||||:|
873 GGGGTGATGGCGGCAATATCGGCATC.....CGGACGTTCA 836
49 nThrgLYPheValArgTYrAspAspIuTYrValSerThSerLeuSer 66
: |||||:|
835 ATCGGGTGTTCAG.....CTCGCG 813
66 euArgSerAlaHisLeuAlaGlyInSerIleLeuSerGlyTYr..... 80
812 AATTCAACSTCGACSTGGGGCGGCACTTCTGCGCGGTTTCANGTAC 763

```


alignment_scores:
 Quality: 89.00 Length: 270
 Ratio: 0.774 Gaps: 14
 Percent Similarity: 42.593 Percent Identity: 22.222

alignment_block:
 US-09-528-682-1 x US-09-540-209B-4736 ..

Align seg 1/1 to: US-09-540-209B-4736 from: 1 to: 3360

```

12  ProProAspGluIleuLysArgSerGlyGlyLeuMetProArgGlyHis 28
    |||  :::::
1348  CCGGACGAGAGTGTCAACGCCATTAACCCATATATAAAGTGGCAGGAA 1397
    ::::: |||  :::::
28  ngIuYrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHis 45
    ::::: |||  |||  :::::
1398  TGAATGGCTTCAACGCTACACCGTTTCAACAGCCCTTTAGCAGAG 1447
    45  laArg.....
    :::::
1448  TCAAGTCTTTAAAGATTAAAGCTGGGATTTCATTCAATTCAACCCGT 1497
    47  .....
1498  TACATCTACGAACCCGCCAATGGGGGTACCTGTTATCAGACAGCTTT 1547
    55  rAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisLeu 72
    ::::: |||  :::::
1548  CAGTGAAGGGGTAAATGTGCAGATGTATCACTCCTCTTCCAA..... 1590
    72  laGlyInSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIle 88
    |||  :::::
1591  ....TTAAGCACAAGTTTCGGTTATAGTCAATTACTCTTAT..... 1629
    89  AlaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValIle 105
    |||  :::::
1630  ...ACATTGGAAAACCTGTTAAACTACAT..... 1656
    105  rProHisProTyrGluGln.....GluValSerAlaLeuGlyIle 120
    |||  :::::
1657  ...CATACATTTCGTCAAGAAACAGATGTATCCGGCTGTAGGC.... 1698
    120  roTyrSerGlnIleTyr...GlyTrpTyrArgValAsnPhe..... 132
    |||  :::::
1699  ...TATCAGGAATTCATTAAGAAATTAAGTACTGATGATCCGCAAGAA 1746
    133  GlyValIleAspGluArgLeuHisArg..... 141
    |||  :::::
1747  GGTTCATGACGAAATCACTGAACAGTTTGATGAAGCAACGAAATGAC 1796
    141  .....
1797  AAGTACCAAGGTGCAACCCAGATTAATGCCACAGTTCGTATTTGGAC 1846
    142  ....AsnArgGlyTyrArgAspArgTyrTyrArgAsnLeuAsnIleAla 156
    |||  :::::
1847  GCGTGAACATATGCAATATAATCCCGGATATTATTATGAGCAAAATTC 1893
    157  ProAlaGluAspGly.....TyrArgLeuAlaI 166
    |||  :::::
1894  ...CGTTATGACGGCTCTCACCTTCCACAAAGATCATCGTGGGGATT 1940
    166  yPheProProAspHisGlnAlaIleTyrArg.....GluGluProTyrIle 181
    |||  :::::
1941  CTTCCCTCTCTTATCAAGTGCATGGGCAATCAAGAAAGAGTTCATCG 1990
    181  tHisIleAlaProGlnGlyCysGlyAsn...SerSerArgThrIleThrGly 196
    ::::: |||  :::::
1991  AAATATACCCCTACCTGCTGATGAACCTCGTAAGTCCGTGATCCTGGGT 2040
    197  AspThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgI 213
  
```

```

2041  AAATAGTACTCTGTGAATTTGGTAACTAC.....GA 2072
    |||  :::::
213  uTyrgInSer 216
    |||  :::::
2073  ATACATCTCG 2082
  
```

seq_name: /cgn2.6/plodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-2852

seq_documentation_block:

; Sequence 2852; Application US/09540209B

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES

; FILE REFERENCE: 2709.1001-001

; CURRENT APPLICATION NUMBER: US/09/540,209B

; NUMBER OF SEQ ID NOS: 10444

; SEQ ID NO 2852

; LENGTH: 3027

; TYPE: DNA

; ORGANISM: B. fragilis

US-09-540-209B-2852

alignment_scores:
 Quality: 88.00 Length: 265
 Ratio: 0.765 Gaps: 13
 Percent Similarity: 43.396 Percent Identity: 21.509

alignment_block:
 US-09-528-682-1 x US-09-540-209B-2852 ..

Align seg 1/1 to: US-09-540-209B-2852 from: 1 to: 3027

```

12  ProProAspGluIleuLysArgSerGlyGlyLeuMetProArgGlyHis 28
    |||  :::::
1348  CCGGACGAGAGTGTCAACGCCAATTAACCCATATATAAAGTGGCAGGAA 1397
    ::::: |||  :::::
28  ngIuYrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHis 45
    ::::: |||  :::::
1398  TGAATGGCTTCAACGCTACACCGTTTCAACAGCCCTTTAGCAGAG 1447
    45  laArg.....
    :::::
1448  TCAAGTCTTTAAAGATTAAAGCTGGGATTTCATTCAATTCAACCCGT 1497
    47  .....
1498  TACATCTACGAACCCGCCAATGGGGGTACCTGTTATCAGACAGCTTT 1547
    55  rAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisLeu 72
    ::::: |||  :::::
1548  CAGTGAAGGGGTAAATGTGCAGATGTATCACTCCTCTTCCAA..... 1590
    72  laGlyInSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIle 88
    |||  :::::
1591  ....TTAAGCACAAGTTTCGGTTATAGTCAATTACTCTTAT..... 1629
    89  AlaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValIle 105
    |||  :::::
1630  ...ACATTGGAAAACCTGTTAAACTACAT..... 1656
    105  rProHisProTyrGluGln.....GluValSerAlaLeuGlyIle 120
    |||  :::::
1657  ...CATACATTTCGTCAAGAAACAGATGTATCCGGCTGTAGGC.... 1698
    120  roTyrSerGlnIleTyr...GlyTrpTyrArgValAsnPhe..... 132
    |||  :::::
1699  ...TATCAGGAATTCATTAAGAAATTAAGTACTGATGATCCGCAAGAA 1746
    133  GlyValIleAspGluArgLeuHisArg..... 141
  
```



```

74 inser.....IleuSerGlyTyrSerThrTyrIleTyrValIle 88
   ||| ..... ||| ..... ||| .....
234 GCAGCCTCATGTGTGGCTGGGGGACG.....CTCATGTGGTG 274
   ||| ..... ||| ..... ||| .....
89 AlaThrAlaProAsnMetPheAsnValAsnAspValIleuGlyAlaTyrSer 105
   ||| ..... ||| ..... ||| .....
275 GCTGGGGGACGCCCCATGTGTGGCTGGGGACAGCCTCATGTGGCAGC 324
   ||| ..... ||| ..... ||| .....
105 rProHisProTyrGluGlnGluValSerAlaLeuGlyGlyIleProTyrS 122
   : ..... : ..... : .....
325 TGGG.....GCAGTGTGGGGGCTT..... 346
   : ..... : ..... : .....
122 erGlnIleTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArg 138
   ||| ..... ||| ..... ||| .....
347 .....GGCGGCTCATGCTGGGAAGTGCATGAGCAGGCCCATC 385
   ||| ..... ||| ..... ||| .....
139 LeuHisArgAsnArgGluTyrTyrArgPArgTyrTyrArg...AsnLeuAs 154
   : ..... : ..... : .....
386 ATTCATTTGGCGAGTGTACTATGAGACCGTACTATCGTAAACATGCCA 435
   ||| ..... ||| ..... ||| .....
154 nIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspH 171
   : ..... : ..... : .....
436 CCCTTACCCCAACCAAGTGTACTACAG.....CCCATGGATG 473
   : ..... : ..... : .....
171 IsGlnAlaTrpArgGluGluProTyrIleHisHisAlaProGlnGlyCys 187
   : ..... : ..... : .....
474 AGTACAGC...AACCAACACACTTGTGTC.....GACTGC 508
   : ..... : ..... : .....
188 GlyAsn.....SerSerArgThrIleThrGlyAspThrCysAsnG 201
   ||| ..... ||| ..... ||| .....
509 GTCAATATCAATCAAGCAGCAGCGTACACCAACCAACCAAGGGGGA 538
   ||| ..... ||| ..... ||| .....
201 uGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerLysv 218
   ||| ..... ||| ..... ||| .....
559 G.....AACTCACCGAGACGACGCTTAAGATGATGACGCCGTGG 599
   ||| ..... ||| ..... ||| .....
218 aLysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 233
   ||| ..... ||| ..... ||| .....
600 TTGAGCAGATGTGTATCACCCAGTACGAGGAATCTCAGGCTTAT 646
   ||| ..... ||| ..... ||| .....
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-252780
seq_documentation_block:
; Sequence 252780, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218.006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198.676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193.483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185.218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167.363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156.358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146.002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 252780
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-252780

```

```

alignment_scores:
  Quality: 85.00      Length: 221
  Ratio: 0.833       Gaps: 10
  Percent Similarity: 46.154   Percent Identity: 19.457
alignment_block:
US-09-528-682-1 x US-10-027-632-252780/rev ..
Align seg 1/1 to reverse of: US-10-027-632-252780 from: 1 to: 2298
51 GlyPheValArgTyrAspAspGlyTyrValSerThr.....SerLe 64
   ||||| ..... ||| ..... ||| .....
934 GGTTTTACATGAATTCATCTGAGTATCTAGAGACTTACTGCTTTTCT 885
   ||||| ..... ||| ..... ||| .....
64 uSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrS 81
   ||||| ..... ||| ..... ||| .....
884 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835
   ||||| ..... ||| ..... ||| .....
81 erThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetPhe..... 95
   : ..... : ..... : .....
834 AG.....GCCACGTGACCACATCTCTTTAGATTC 806
   : ..... : ..... : .....
96 .....AsnValAs 98
   : ..... : ..... : .....
805 TCTCCAAAGAAATATTGAAGAGATCGCTCATCCCATGCTATACAAATATC 756
   : ..... : ..... : .....
98 nAspValIleuGlyValTyrSerProHisProTyrGluGlnGluValSerA 115
   : ..... : ..... : .....
755 AGATCATTTATCTGTCTGACCTCTCA..... 731
   : ..... : ..... : .....
115 lAlaGlyGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsn 131
   ||| ..... ||| ..... ||| .....
730 .....CCCCAAACTAAACTCTCAAGTTTCTGAATTTAA 695
   : ..... : ..... : .....
132 PheGlyValIleAspGluArgGlnHisArgAsnArgGluTyrTyrArgPArg 148
   : ..... : ..... : .....
694 AAAAATTCATTTACCAACACATGCAAT...CATAGTATCATAGGACAA 648
   : ..... : ..... : .....
148 gTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuA 165
   : ..... : ..... : .....
647 TGAGATTATGAAGATGCATGCCATCTCTCAGAGCGCTCATATCTATTT. 599
   : ..... : ..... : .....
165 lAglyPheProProAspHisGlnAlaTrpArgGluGluProTyrIleHis 181
   ||| ..... ||| ..... ||| .....
598 .....TGCAGAGAACACATCATCATGCAC 575
   : ..... : ..... : .....
182 ...HisAlaProGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAs 197
   ||| ..... ||| ..... ||| .....
574 GTCCACACA.....TGCACAAATTTGTTCAAGACAAACCAAGA 534
   : ..... : ..... : .....
197 PThr..... 198
   : ..... : ..... : .....
533 GACCCATCAAGCGTTATGTCCTTTTCCAAATGATGAGCAGGAT 484
   : ..... : ..... : .....
199 .....CysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArg 212
   ||| ..... ||| ..... ||| .....
483 AGCTGATTTGTTTAAACGAGGAGAACTTACAC...CTGTGCTGACA 437
   : ..... : ..... : .....
213 GluTyrGlnSerLysValLysArgGlnIlePheSerAspTyrGlnSerG 229
   : ..... : ..... : .....
436 GGAGGAAGACAGCAGTAAACAAAGAGACTATGATAGCTAGTGTACA 387
   : ..... : ..... : .....
229 uValAspIleTyr 233
   ||| ..... ||| ..... ||| .....
386 GCTCAATTTGTTAT 374
   : ..... : ..... : .....
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-919-002-8511
seq_documentation_block:
; Sequence 8511, Application US/09919002
; GENERAL INFORMATION:

```



```
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/919,002
; CURRENT FILING DATE: 2001-07-30
; PRIOR FILING DATE: FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
; PRIOR FILING DATE: FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8511
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-002-8511
```

```
alignment_scores:
  Quality: 84.00      Length: 168
  Ratio: 0.977       Gaps: 10
  Percent Similarity: 51.190   Percent Identity: 23.810
```

alignment_block:

US-09-528-682-1 x US-09-919-002-8511/rev ..

Align seg 1/1 to reverse of: US-09-919-002-8511 from: 1 to: 858

```
53 ValArgTyrAspAspGlyTyrValSerThrSerLeuArgSerAl 69
|||||:|||||:|||||:|||||:|||||:
652 GTTCGGTTGATGCGCTGATATACGATAAAACACATATATCTGCTAA.. 605
69 aHsLeuAlaGlyInSerIleLeuSerGlyTyr..... 80
|||||:|||||:|||||:|||||:|||||:
604 .....GGGGACAGCTCTTGTATGTTCTATAGAGCCGTCGACCC 565
81 ..SerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsn 96
|||||:|||||:|||||:|||||:|||||:
564 AAGTGAATTTACAGGTACATA.....AGATTCTTCCCT 530
97 ValAsnAspValLeuGlyValTyrSerProHisProTyrGluGlnGlu 113
|||||:|||||:|||||:|||||:|||||:
529 GATGGCCATGTGATGATGTGACAAACCCTGAAAGCCCTCAGTCCATTGT 480
113 lSerAlaLeu...Gly GlyIleProTyrSerGlnIleTyrGlyTyrTyr 128
|||||:|||||:|||||:|||||:|||||:
479 TCCACGTTTAAAGTAAGATACAGAGCTGATGCAATTCTCTGCTGCTC 430
129 ArgValAsn.....PheGlyVal 134
|||||:|||||:|||||:|||||:
429 ACATTCCTGTGTCACAGACACAGACATCAGACCAAAAGATTTGCTGT 380
134 lIleAspGluArgLeu...HisArgAsnArgGlyTyrArgAspArgTyr 150
|||||:|||||:|||||:|||||:|||||:
379 AATTAACTAAGAAAAAGAGAAAAACCACTTGACTATATAATACAGATTT 330
150 YrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArg.....Leu 164
|||||:|||||:|||||:|||||:|||||:
329 TTGCTGCGTCCCTGTACAGAGACAGATCAGAGTTTTCATGTGGGCTTA 280
165 AlaAlaGlyPheProAspHisGlnAlaTyrArgGluGluProTyrPheH 181
|||||:|||||:|||||:|||||:|||||:
279 CAGCTATGTTCCAGTGGTCAACAGAGTTTCAACAACCTCATCTGGATACA 230
181 SHSAlaProGlnGlyscIyAsnSerSerArgTyrIleThrGlyAspT 198
|||||:|||||:|||||:|||||:|||||:
229 TCATTCT.....TGTCACATTACTTACAATCA...ACTGGTGAGA 192
198 hr 198
||
191 CT 190
```

seq_name: /cgn2.6/plodata/1/pna/US09_NEW.COMB.seq:US-09-919-002-6157

seq_documentation_block:

; Sequence 6157, Application US/09919002

; GENERAL INFORMATION:

; APPLICANT: Leshkowitz, Dena

; APPLICANT: Liu, Jin

; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA

; FILE REFERENCE: 20411-752CON1

; CURRENT APPLICATION NUMBER: US/09/919,002

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922

; PRIOR FILING DATE: FILING DATE: 1999-07-22

; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341

; PRIOR FILING DATE: FILING DATE: 1998-02-13

; NUMBER OF SEQ ID NOS: 13203

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 6157

; LENGTH: 3617

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-919-002-6157

alignment_scores:

```
Quality: 83.00      Length: 228
Ratio: 0.761       Gaps: 16
Percent Similarity: 47.807   Percent Identity: 22.368
```

alignment_block:

US-09-528-682-1 x US-09-919-002-6157 ..

Align seg 1/1 to: US-09-919-002-6157 from: 1 to: 3617

```
40 AsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArgTyrAs 56
|||||:|||||:|||||:|||||:|||||:
2132 CACCTTACTGATACATGATGATGCTAGTACGAGAAAAACCTGCAAAATATGA 2181
56 pAsp.....GlyTyrValSer...ThrSerLeuSerLeuArgSerA 69
|||||:|||||:|||||:|||||:|||||:
2182 AGAGTCGACGAGTGTCTTTATCTTGTCTTCACAGCCCTTCTAATCAACAGA 2231
69 lAsnIleuAlaGlyInSerIleLeuSerGlyTyrSerThrTyrTyrIle 85
|||||:|||||:|||||:|||||:|||||:
2232 TGATTCTTCTGCGAGAGAG..CTCTCCAAATGTGAACATGTGTAC... 2275
86 TyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspValLeuG 102
|||||:|||||:|||||:|||||:|||||:
2276 .....AAAGTTTAAAC..... 2287
102 yAlaTyrSerProHisProTyrGluGlnGluValSerAlaLeuGlyTyr 119
|||||:|||||:|||||:|||||:|||||:
2288 ....CACAGCCAAATCTTCCAAACACAGAGAAATGAGATTTGGAGGA 2333
119 lProTyr.....SerGlnIleTyrGlyTyr..... 128
|||||:|||||:|||||:|||||:|||||:
2334 AACCTTCAATGTGAGAGATGTGACAGCATCTTCAAGTGTCTCAGAC 2383
129 .....ArgValAsnPheGlyValIleIleAspGluArgLeuH 140
|||||:|||||:|||||:|||||:|||||:
2384 CTTACTTAACATTAAGAGATTTCACTGCT.....GAGAAACCAATA 2424
140 sArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAla 156
|||||:|||||:|||||:|||||:|||||:
2425 CAATGTGACGAATGTGGGAAAGCCTATATACACAGTCTCACACCTCAGTG 2474
157 .....ProAlaGluAspGlyTyrArg..... 163
|||||:|||||:|||||:|||||:|||||:
2475 AACACAGAGAGATTACACCGGAGAGAGAAACCTTACCAATGTGAAGATGT 2524
164 .....LeuAlaGlyPheProAspHis 171
```



```

2525 GCGAGGCTTTGAGAACTTGTGCTCAAGCCTTTCTACATAGAGAACTCA 2574
171 SGlnAlaTrpArgGluInuProTPrIleHsHsAlaProGlnGlyCysG 188
2575 TTCT.....GAAGAAAAACCC.....TACACCTGTGAAGAAATGTG 2609
188 Lysn.....SerSerArgThrIle 194
2610 GCAACATCTTTAAGCACTATACAGACCTCACTAAGCATAGAAAAACCAT 2659
195 ThrGly.....AspThrCysasnGluInu...ThrG 204
2660 ACTGAGAGAGAGAGCCCTACAAATGTGACGATGTGAGAAAAAATTACCA 2709
204 nasLeuSerThrIle.TyrLeuArgGluTyr 214
2710 GTCCTCCACCTTATGTGTACATAGAGAAATTC 2741

```

seq_name: /cgn2_6/pdata/1/pna/US60_NEW_COMB.seq:US-60-360-039-32681

```

seq_documentation_block:
; Sequence 32681, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; NUMBER OF SEQ ID NOS: 2002-02-21
; SEQ ID NO 32681
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Chloroflexus aurantiacus
US-60-360-039-32681

```

alignment_scores: Quality: 82.50 Length: 219 Ratio: 0.764 Gaps: 12 Percent Similarity: 49.315 Percent Identity: 26.027

alignment_block:

US-09-528-682-1 x US-60-360-039-32681 ..

Align seg 1/1 to: US-60-360-039-32681 from: 1 to: 1107

```

3 AspArgLeuTyrArgAlaAspSerArgProProAspGluIleLysArgSe 19
141 GACCGCATTTGGCGATGACGAGCCGACCACTGACAGCAGCCGGATATC 190
19 rGlyGlyLeuMetProArgGlyHisn.....G 29
191 G.....CGTTCACAGCAGCACTGTGCCATCAGGTGAGACAG 225
29 LutyPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAla 45
226 AGCAAAATGCTTCGGCGAAGTGAACCGACTGTACAGCTGGATGATGAGA 275
46 Arg.....GlyThrGln..ThrGlyPheValA 54
276 CGGAAATGGAACGCCGTGATGCTGCTGCGCCCTGGTGGTGGCGGTGCG 325
54 rGlyTrpAspArgGlyTyrValSerThrSerLeuSerLeuArgSerAlaHis 70
326 GTGACCTGGCAGGGTTTGGCGCAGCAGCAATC..... 357
71 LeuAlaGlyInserIleLeuSerGlyTyrSerThrTyrIleTyrVal 87

```

```

358 TTGGCGCGGTGGGGCTTCGTCAGATACCGAGCAGCGCTCCGGGATGCT 407
87 LLeAlaThrAlaProAsnMetPheAsnValAsnAspValIleGly.... 102
408 TGATGCTGGCGGTAGCTGTAAACCGGCATCAACACCGCTGGGCAAA 457
103 .....ValTyrSerProHisProTyrGluInGlnValSer 114
458 ACCTGATCGGAGCCTTTCATCAACACCGGTGGTGGTGGCGACACCAAC 507
115 AlaLeuGlyGlyIleProTyrSerGlnIle...TyrGlyTyrTyrArgVa 130
508 CTGTGGCGACACTACCGCGCTCGGAGCTACGCGCCGCTGGCGGAAGT 557
130 L...AspPheGlyValIle...AspGluArgLeuHis...ArgAsnArgG 144
558 GATCAACACGCTGTCATTCGCGATGCGCGAATGTGTGGTGGCGGAA 607
144 uTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAsp 161
608 CACTGCG.....CGCAACCCAAACGCTGGCACGCGGACCAATCCG 645
161 LutyArgLeuAlaGlyPheProProAspHisGlnAlaTrpArgGluInu 177
646 GCTACCTGGGATGGCGACCGATCCGCGCTCACCGGTGAACAGACCA... 692
178 ProTPrIleHisAlaProGlnGlyCysGlyAsnSerSerArgThrI 194
693 .....CATCAATCGCGCGCTGTGGTGGTGAAGTGCAGCGTGTGAT 733
194 eThr 195
734 CACG 737

```

seq_name: /cgn2_6/pdata/1/pna/US60_NEW_COMB.seq:US-60-360-039-32252

```

seq_documentation_block:
; Sequence 32252, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; NUMBER OF SEQ ID NOS: 2002-02-21
; SEQ ID NO 32252
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Ralstonia metallidurans
US-60-360-039-32252

```

alignment_scores: Quality: 82.50 Length: 313 Ratio: 0.640 Gaps: 16 Percent Similarity: 41.214 Percent Identity: 20.447

alignment_block:

US-09-528-682-1 x US-60-360-039-32252 ..

Align seg 1/1 to: US-60-360-039-32252 from: 1 to: 1524

```

8 AlaAspSerArgProProAspGluIleLysArgSerGly..... 20
415 GCGGACAGCGCTCGATGAGAGAGCGCTTACCGTCAACGCGCAGCACTGGC 464
21 GlyLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGln 37

```



```

465  |||:..... ||| ... ||| :|||:.....
    GGGGATTTTCTACACCGGAGGACACCGCGTCCCGAAGGGGGATGC 514
37  eAsnIleasnLeuTyraSpHisAlaArgGlyThrGlnThrGlyPheVal 53
515  TCAGC..... |||:|||||:.....
    CATGGAGACCTGATGCTCTGGCATGCAAC 549
54  ArgTyraSpAspGlyTyraValSerThrSerLeu..... 64
550  CAGCGGATGTCGGGACGCGTGGATGAAATCCATCGATGATGATGATGAT 599
65  ..... SerLeuArgSerAlaHisLeu 72
600  GCCGATGTTCCACCTGCGCATGCTTCCGCTGCATCAACCGCTGTCCTCG 649
72  IagIlyGlnSer..... Ileu 77
650  TCGGTGGCAGCAGATGGCGGATATCCGCTTCGATCCTGGCCGCGATGATG 699
78  SerGlyTyraSerThrTyraIleTyraValIleAlaThrAlaProAsnMe 94
700  GAGGGGATTTGGCAGACAGGATGACGGCGCTCGCTGGGGCCGACCAT 749
94  tPheAsnValAsnAspValIleuGlyValTyraSerProHisPro..... 108
750  GATCCAGATGGGG..... CTGGACTGGATGACCGCGCATCCGGGGCGCG 793
109  ..TyrlGlnGlnValSerAlaLeuGlyIleProTyra..... 121
794  CGCGCGAGCTGGACCTGTGCTGGTGGACAGCGCTGGCTATGGCGCTCG 843
121  ..... 121
844  CCGATGACCCCTGGCGCTCTGGAGGGACACGAGACACTTCCAGCGGT 893
122  ..... SerGlnIleTyrlGlyTrp..... 127
894  GCGCTTTTCGAGGGCTATGGCATGACCGAGCTGGACCGGTGGCCACAA 943
127  ..... 127
944  TGCCTGGCCCGGAGTATGACAGCGCGGCGGCGGACGCGCAAGATG 993
128  ..... TyraIyValAsnPheGlyValIleAspGly 137
994  TATTCAGTCGCGCGCGCGCGCATACGCTGAGGTCAAGTGTGCTGACTTC 1043
137  uArg..LeuHisArgAsnArg..GluTyraArgAspArgTyraTyraArgAsnLe 153
1044  CAGGGGCAACAGGTCCGCGCGGTACCGTGGGCGAGATCATCTGACTG 1093
153  uAsn..... IleAlaProAlaGlu..... 159
1094  GCCCGAATGTATGCTCGGCTACTGGAACCGCGGACGCTACCGCGGAG 1143
160  ..... AspGlyTyraArg..... 163
1144  GCGATCCGCAATGGCTGATGATACCGGCGAGCGTGGTACATGACAA 1193
164  ..... LeuAlaGlyPheProPro..... AspHisGlnAl 173
1194  GGACGCGTTCCTGTCTGTCGTGACCGCGCTCAAGACATGATCATCACCG 1243
173  aTrrArgGluGluProTrpIleHisHisAlaProGlnGlyCysGlyAsn 190
1244  GTGGCGAGAACGT...CTATTCGCGCGGAGTGGAGAAAGGCGCTGGCGAGT 1290
190  erSer..... ArgThrIleThrGlyAspThrCys 199
1291  CATCGGCGCGGTGGCGAGTGTCTGTATGCGGTGT 1325
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-1507

```

```

seq_documentation_block:
; Sequence 1507, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1507
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: B. fragilis
; US-09-540-209B-1507

alignment_scores:
    Quality: 82.50      Length: 219
    Ratio: 0.851       Gaps: 18
    Percent Similarity: 44.292    Percent Identity: 26.027

alignment_block:
US-09-528-682-1 x US-09-540-209B-1507 ..
Align seg 1/1 to: US-09-540-209B-1507 from: 1 to: 3000

17  LysArgSerGlyGlyLeuMetProArgLysAsnGlyTrpPheAsp 33
609  AAGGGGTTTGGACGGTATCAATCCCGAGACATCCG.....TTCGATGC 652
33  gGlyThrGlnMetAsnIleasnLeuTyraSpHisAlaArgGlyThrGln 50
653  AGGTTTCAAGAAGATCATC...GTCCACTGCCATATACGG..... 689
50  hrGlyPheValArgTyraSpAspGlyTyraValSerThrSerLeu 66
690  ....TTCAGGGGCTTCGAACGGGT..... 710
67  ArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyraSerThr 83
711  CGTCTGATCATC.....TACCAAGATTG 733
83  r.....TyrlTyraValIleAlaThrAla..Pro 92
734  GTAGGCGCGGTGTGCGAATATATGTTTCAGCAGCATGAGTGGGATGCTCC 783
93  AsnMetPheAsnValAsnAspValIleuGlyValTyraSerProHisPro 109
784  AMGTATATTAAGATGATGATATCCCTTGGAGCATACGAT.....TA 824
109  rGlnGlnGlnValSerAlaLeuGlyGlyIleProTyraSerGlnIleTyrg 126
825  CGCTCTGGCGCTCAAGGAGGTCAAGGATATTGATTCTCA..... 864
126  LyrTrpTyraValAsnPheGlyValIleAspGluArgLeuHisArgAsn 142
865  .....AAC 867
143  ArgGluTyraArgAspArgTyraTyraArgAsnLeuAsnIleAlaProAla 159
868  GAAGAGATGCAATCC.....TATCAGAACCG.....AA 895
159  uAspGlyTyraArgLeuAlaGly.....PhePro.....Pro 170
896  CCGGGGCGATGCGACTGGCAGATGAGATTTTCCGTACGGGGAATCACCAG 945
170  spHisGlnAla.....Trp...ArgGluGluProTrpIleHis 181
946  AATTACAAAGTTGGCTCTTCCAAATGTGAAGAGAAAGCCCAATATTATCAT 995
182  .....HisAlaProGlnGlyCysGlyAsn.....SerSer 192

```


996 TTCGCCCACTACATGAGCCAGGAAGGTGTGTCATGATCGAAGAACG 1045
192 gTthrlhrlglsprhrcysasnlgltlthrclnasnleuserthr 209
1046 AGCGTATACAGC.....GAAGGCGCAATCTTCTCCACAG 1080
209 leTyr 210
1081 CTTCAC 1085
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-209B-2633
seq_documentation_block:
; Sequence 2633, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 2633
; LENGTH: 723
; TYPE: DNA
; ORGANISM: B. fragilis
US-09-540-209B-2633

alignment_scores:

Quality:	82.00	Length:	271
Ratio:	0.678	Gaps:	15
Percent Similarity:	44.649	Percent Identity:	20.295

alignment_block:

US-09-528-682-1 x US-09-540-209B-2633 ..

Align seg 1/1 to: US-09-540-209B-2633 from: 1 to: 723

```
14 Aspgluilelvsargserglyleuemetproarglyhls..... 27
   |||||
16 GAAAGATATAAATCTGCGTCATCTTTATAGATAAACAACCTCGGCT 65
   |||||
28 .....AsngluTyrpheasprarglythrlnmetasnlea 40
   |||||
66 ACACACAATGAAGAACAAACGCTACTCAGCGCGAGAGAGCTGTCC...A 112
   |||||
40 snleuTyrAspHisAlaIArglyThrlnThrlnGlyPheValArgTyrAsp 56
   |||||
113 ATACTTTAGCCATGCGCAGACACCTGCTGGCATCACCGCA..... 156
   |||||
57 AspglyTyrValSerThrSerleuserleuargseralanhls..... 70
   |||||
157 ...GGATATTTCTTTTGAAAGAGCGCTCGCGCAACCTCACCCCTATTG 203
   |||||
71 .....LeuAlaGlyGlnSerlleuserg 79
   |||||
204 GCGAAGACAGCTGTGACTGCGCTTATCTGTGTGGG.....ATGCTGCAT 247
   |||||
79 lYTyR.....SerThTyTyR..... 84
   |||||
248 CGTACATCATGTTGACCTGTATCAGCGCTCCAGCCCGGCAAGCGGAAA 297
   |||||
85 .....lTeyrValIleAlaThrAl 91
   |||||
298 GAACCTTGGCGCAATTTGACACAGCGGCATTTTACCTG..... 336
   |||||
91 aProAsnMetPheAsnValAsnAspValleuGlyValTyrSerPro..Hl 107
   |||||
337 .....CACATGACAGGAGACTTACACTCGGTCA 364
   |||||
107 sProTyrGlnGlnIuValSerAla.....LeuG 117
   |||||
```

```
365 CCCT.....GTAGTGTACAGCATGCCGGGAGTGGCTGG 402
117 lYgIlyleProTyrSerGlnleTyrGlyTyrPyrArgValAsnRheGly 133
   |||||
403 GGAATCTTCACATTGTGTAGCGCTGCGCTATTGTAGGATTCATATAGC 452
   |||||
134 ValIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTy 150
   |||||
453 CT.....CAAAAACSTGAAGAACAAGCAATCTCGACAGCATCTGT 496
   |||||
150 rArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyP 167
   |||||
497 TTGTAGGAGTGGCTGTGCAATCCGTGGGCGCCCAACCACTATATAGAC 546
   |||||
547 TCCCTGTCGCCATCGGGGCGCTCCGCTTGTGTGCTGTCTGTGAGAG 596
   |||||
177 GluProTyrPheHisAlaProGlnGlySerGlyAsnSerSerArgTh 193
   |||||
597 AGGACCTTCGTATATCAT.....GGAGCGCTGT 625
   |||||
193 rIleThrGlyAspThrCysAsnGluGluThrGlnAsnleuserthrIleT 210
   |||||
626 TTTACTC.....GTGCGAAGACCTTAT 648
   |||||
210 YrLeuArgGluTyrGlnSerLysValLysArgGlnIlePheSerArgPyr 226
   |||||
649 ATGACAGCGCGTTTCCACCTGTTGTGTAGCGGAGAGCATAGCCATAT 698
   |||||
227 GlnSerGluVal 230
   |||||
699 CATAGCCATCTG 710
```

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-27746

seq_documentation_block:

; Sequence 27746, Application US/60360039

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Chen, Xianfeng

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)A

; CURRENT APPLICATION NUMBER: US/60/360.039

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 27746

; LENGTH: 3291

; TYPE: DNA

; ORGANISM: Neurospora crassa

US-60-360-039-27746

alignment_scores:

Quality:	82.00	Length:	135
Ratio:	1.139	Gaps:	7
Percent Similarity:	53.333	Percent Identity:	25.185

alignment_block:

US-09-528-682-1 x US-60-360-039-27746 ..

Align seg 1/1 to: US-60-360-039-27746 from: 1 to: 3291

```
56 AspaArglyTyrValSerThrSerleuserleuargseralanhleuAl 72
   |||||
2017 GACATGGCGACAGATCGAGGCGAGATTCGACTGAGTTTAACTAGC 2066
   |||||
72 nGlylnSerlleuserGlyTyrSerThTyTyrIleGlyValIleAla 89
   |||||
```



```

2067 GTTATATCCCTG.....GGCTGATGAC...TCATGTTCTGTAATAG 2107
89 lathralproasmetpheasnvalasnpval..... 100
||||| :||| :|||
2108 CAACAGCAACACCGGCTTCATTATACCCAGCGTTGGCAGCAGCAATGTA 2157
101 ..leuglyvaltyrserprohis..Protyrgluglnlvalseral 115
:|:|:| :|||:|:| :|||:|:|
2158 TTGTTCTATTTGATGTCGTTACCCATCCGATCGCAGAGCTGGCTGCG 2207
115 aleuglyglileprotyrserlsoletyrglytrpyrtyrvalasnp 132
||||| :||| :|||
2208 TCTCCAGACGGCTCCCTGCTGGCCCTATATGTC.....AATT 2245
132 hseglyvalleasprgluatrgleuhsisargasnargslutyrg...Asp 147
||||| :||| :|||
2246 TTGGCAGCGTTCTCGAAGCGTTGACAGCGTACGTCCTCCGCGCAGAG 2295
148 Arglytyrtyrargasnleuhsnlealproalaglualspglytyrargle 164
:|:|:|:| :|||:|:| :|||:|:|
2296 CCACATTTCCAGAACCGCATTCATCCGACCGCATGACTTTCAGAAAT 2345
164 valagltyrheproproasphslnalatrparglugluprotprileh 181
2346 G.....GATCACAATATATCTGCTGCTGCTGCTGCTGCTGCTGCTG 2380
181 lshis 182
:|:|:|:|
2381 AATAC 2385

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-28417
seq_documentation_block:
: Sequence 28417, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkley, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-10152052JA
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 28417
: LENGTH: 2028
: TYPE: DNA
: ORGANISM: Burkholderia fungorum
US-60-360-039-28417

alignment_scores:
Quality: 81.50 Length: 303
Ratio: 0.709 Gaps: 11
Percent Similarity: 37.954 Percent Identity: 19.142

alignment_block:
US-09-528-682-1 x US-60-360-039-28417 ..
Align seg 1/1 to: US-60-360-039-28417 from: 1 to: 2028

9 Asperargproproasp.....gluul 16
|||:|:|:|:|
253 GACGGCAACCGCAGCGACCGACCGAGTTGCTGACGATATCGAACA 302
16 elysahsargglyglyleumetproarglyhisasnlytyr..... 30
|||:|:|:|:|
303 TTTCGCCGATGGGAGAGACACCGGCGCAGTCCGAGTCCGATGA 352
31 .....Pheasprgltyrthrlnmet 37
:|:|:|:|

```

```

353 CCACGGGCTGCGAGCAGCAGCAGCGGCGCTGGGCAAGGGCTCGGCAT 402
38 Asnileasleuatyasphslnalatrarglythrlntrgltyrvalar 54
:|:|:|:|:|:|:|
403 AGCGTTGGCATGGCATGGCGGCGCTGTGTAAGAGACATTTCAACA 452
54 gtyrasp.....Asp 58
:|:|:|:|
453 GCCCGATGCGCGCTGTTGACATATCGGGTCTACGCGTTGTGCGCGAG 502
58 lutyvalserthrserleuaserleuargseralhisleuagllyln 74
|||:|:|:|:|:|:|
503 GCGACATATGGAAGACATCTGCGACGAGGGGCTTCGCTGCGGCGAC 552
75 Serileuser..... 78
|||||
553 CTGAAGCTGTCCAAATCTGATCTGATCTACGACAGCAACCGCGTGCAT 602
78 ..... 78
603 TGAAGGCCACACGACCTCGCTTATAGCGAGCGTGGAAACCGCTTTC 652
79 ..glytyrserthrttyrlyletyrval..... 87
|||||:|:|:|:|
653 GCGGTACACTGGCACACGCTGCTGATGTAGCGATGCAACGACCGCGC 702
88 .....lealatrlnalproas 93
|||||
703 GCGCTGAAGCGCGCATTCGTGCAAGCAAAAGCATATACGACAGCGCAC 752
93 metherasnaValasnpvalleuglyvaltyrserprohisprotyrg 110
:|:|:|:|:|:|:|
753 GCTGATCTGTGGCACACATCATCATGCTGGGCGCGCGCAT.....A 796
110 luglnlvalseralaleuuglyleprotyr..... 121
:|:|:|:|:|:|:|
797 AGCAGGACACCTCGGCGCGCACGCGGACCGCTCGCGTGAAGAACTG 846
122 .....Serlnlletyrlytyrtyrtyrvalasnpheglyval 135
:|:|:|:|:|:|:|
847 GCGCTGCCAAGAAAGCGTACGCTGGCGCGGAGTAATCTTCTATGT 896
135 eapgluatrgleuhsisarg..... 141
:|:|:|:|:|:|:|
897 GCCGACGCTGTGCAAGACGCTTCGCGCGCGCATCGCGCGCGGCA 946
142 .....Asnarglutyrtargasprtyr..... 149
:|:|:|:|:|:|:|
947 AGCGCGCGCGGAGGATGGCAGCGGAGATACGACGCTACAGCAAGAA 996
150 tyrtargasnleu.....Asnlealproalaglualspglytyrargle 164
|||:|:|:|:|
997 TACCCGAACCTGCGCGTGAATTCGCGAGATCGAGACGACGAAATGCC 1046
164 valagltyrheproproasphslnalatrparglugluprotprileh 181
|||:|:|:|:|
1047 GCCCGTGGGACAGCAGCATTCCTCCATCTGACGCGGACCG..... 1089
181 lshisalaproglnglycysgluasnserserargthllethglyasp 197
|||:|:|:|:|
1090 .....AAAGCGTTGCTGACGCGAATCTGCGGCAAG 1122
198 ThrCysasn 200
:|:|:|:|
1123 GTGCTCAT 1131

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-31176
seq_documentation_block:
: Sequence 31176, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng

```



```

; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 31176
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
US-60-360-039-31176

```

```

alignment_scores:
  Quality: 81.50      Length: 303
  Ratio: 0.709       Gaps: 11
  Percent Similarity: 37.954  Percent Identity: 19.142

```

alignment_block:

```

US-09-528-682-1 x US-60-360-039-31176  ..
Align seg 1/1  to: US-60-360-039-31176  from: 1  to: 2040

```

```

9 AspSerArgProProasp.....Glu11 16
  |||:|||||  |||
250 GAGCGGACCCGCGAGCCGCGAGTTTCCTCGACATATCGACMA 299
  |||:|||||  |||
16 elysargserglyleuwerproarglyhisaanglytyr..... 30
  |||:|||||  |||
300 TTTCCCGCATGGCGACGACAGACCGCGCCCGCGAGTACCGCATCA 349
  |||:|||||  |||
31 .....Pheasparglythriginet 37
  |||:|||||  |||
350 CCACGCGCGTCGAGACACACGCGCGCGCTGGCGAAGGCTCGGCAAT 399
  |||:|||||  |||
38 AsnIlaasnleutyraasphialaarglythrglntheglyphevalar 54
  |||:|||||  |||
400 AGCGTGGGATGGCGATGGCGCGCGCTGTACGAGAGCCATTCAACA 449
  |||:|||||  |||
54 gtyrasp.....Aspg 58
  |||:|||||  |||
450 GCGCGATGGCGCGCTGTGACATATCGGCTCTACGCGCTGTGCGGCG 499
  |||:|||||  |||
58 ltyrvalaserthrserleuaserleuargseralaha1sleualaglyln 74
  |||:|||||  |||
500 GCGACATGATGGAAGCATCTCGCACGAGCGCGCTTCGCTCGCGCGCAC 549
  |||:|||||  |||
75 Serlleuaser..... 78
  |||:|||||  |||
550 CTGACGCTGTCCAACTGATCTGATCTAGACAGCAACCGCGTGACGAT 599
  |||:|||||  |||
78 ..... 78
  |||:|||||  |||
600 TGAAGCCACACGCGACTGCTTATAGCAGACGAGTGGAAACCGCTTC 649
  |||:|||||  |||
79 ..glytyrserthrtyrtyrleuval..... 87
  |||:|||||  |||
650 GCGGTAACAACCTGGCACACGCTGATGTAGCATGCAAAAGCAGCGCGC 699
  |||:|||||  |||
88 .....Ilealathrlapras 93
  |||:|||||  |||
700 GCGCTGGAAACGGCATTCGTGAAGCGAAAGCATTAACGACAGCGCGAC 749
  |||:|||||  |||
93 nmetpheasnvalasnaspvalleuaglyvaltyrserproh1sleuarg 110
  |||:|||||  |||
750 GCTGATCGTGTGACAGCATCATCGCTGGGCGCGCGCAT.....A 793
  |||:|||||  |||
110 luginluvalseralaleuaglylyleprotyr..... 121
  |||:|||||  |||
794 AGCAGGACACCTCGCGCGCGCGAGCGCGCTCGGCGTGAAGAAGATG 843

```

```

122 .....SerGlnleutyrglytrptryargvalasnphleglyval11 135
  |||:|||||  |||
844 GCGCTGGCGCGAGAAAGCGTACGGCTGGCGCGAGATTAATCTTCTATGT 893
  |||:|||||  |||
135 eaargluargleuhs1arg..... 141
  |||:|||||  |||
894 GCGGAGCGGTGTCCACGAAACGCTTCGGCGCGGCGCATTCGCGCGCGCA 943
  |||:|||||  |||
142 .....Asnargglytyrargspargtyr..... 149
  |||:|||||  |||
944 AGCGCGCGCGCGAGATTCGACGCGAGATGACGACGCTACACAGAAAG 993
  |||:|||||  |||
150 TyraArgasnleu.....AsnIlealaproalagluaspelytyrargle 164
  |||:|||||  |||
994 TACCGCGAATCGCGCGGTGAATTCGCGCGAGATGACGACGCAATAATGCC 1043
  |||:|||||  |||
164 ualaglypheproproasphiasglnalatrparargglugluprotrp1en 181
  |||:|||||  |||
1044 GCGCGGTGGACAGCAGCATTCGCCACTTCGACGCGGACCG..... 1086
  |||:|||||  |||
181 lsh1salaoproglnglycysglyasnseraragthrllethrglyasp 197
  |||:|||||  |||
1087 .....AAAGCGGTGGCTGCACGCGAATGCTCGGCGCAAG 1119
  |||:|||||  |||
198 Thrcysasn 200
  |||:|||||  |||
1120 GTGCTCAAT 1128

```

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-42281

```

seq_documentation_block:
; Sequence 42281, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 42281
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Halobacterium sp. NRC-1
US-60-360-039-42281

```

```

alignment_scores:
  Quality: 81.00      Length: 127
  Ratio: 1.397       Gaps: 5
  Percent Similarity: 45.669  Percent Identity: 24.409

```

alignment_block:

```

US-09-528-682-1 x US-60-360-039-42281  ..
Align seg 1/1  to: US-60-360-039-42281  from: 1  to: 2004

```

```

98 AsnAspvalleuaglyvaltyrserproh1sleuarg 114
  |||:|||||  |||
418 GACGACATCTCAACGCTGACATGCGCGATGCGCGCGATCGCGACGC 467
  |||:|||||  |||
114 lalaleuagly.....glyileprotyrsergln1leuarglyt 127
  |||:|||||  |||
468 GATGCTGCGCTGGCGCGCGCTCGCGCGCGCACACTCGGTGTGTCGCG 517
  |||:|||||  |||
127 rpyrargvalasnphleglyval1leaspluargleuhs1argasnarg 143
  |||:|||||  |||
518 GCTTCGCGCG..... 528

```



```

1466 CTACCTGACTTCAGCTATCTACTACAGCGCTACAGTAACCAAAACAGCATG 1417
87 Val11le1aTh1alPProAsmEthEasnuValasnAp..... 99
|||||: : : : :
1416 GTACTGTATTCAAAACAGAGATTGTAGATCAAGAAGAACAGAGCCCT 1367
99 ..... 99
1366 CAGAAATAAGCGGCATATCTCAACAATCTGATCTTTGACAAACCTGAG 1317
100 ..... ValEugly 102
|||||||
1316 AAAAACAAGCAATGGGGAAGGATTCCCTATTTAATAATGATGCTGGGA 1267
103 ValTyserProHisProtYrGLngInguValIsertAlaleuGLyL11 119
|||||: : : : :
1266 AACTGGCTAGCCATATGTACAAAGCTGAAC...GGATTCCT 1226
119 eProtYSerGlnIleTYrGLYTPYrFArgValasnPhEgylVal1lea 136
|||||: : : : :
1225 TCCTTATACC.....TTATACAAAAATCAATTC..AAGATGG 1191
136 spGLuArGLeu.....HisArGsArGLuTYrArg 146
|| ||||| | : : : :
1190 ATTAAGAAGCTTTAAAGCTTAGACCTTAAACATMAAAACCTAGAAGAAA 1141
147 AsParGTyTYrFArgAsnLeuasnIlelaProAlaIGluAspGLYTYrAr 163
|||||: : : : :
1140 CCTAGGCATTAC..... 1129
163 glEuAlaGLYPheProProAsPHisGLnaLTPrArgGLuGLuProTPri 180
|||||: : : : :
1128 .....CATTCAGGATCATGGCATTGGCAGGAGCA 1101
180 leHis...HisLabroGInGLYCysGlyAsnSerSerArGrThr..Ile 194
|||||: : : : :
1100 TTCATGTGCCAAACACCAAACCAATGGCACAAACCAAGCCAATTTGACAA 1051
195 ThrGlyAspThrCysAsnGLuGLUthrGlnasnLeuSerThrIleTYr.L 211
|||||: : : : :
1050 ANGGGATCTAATTAAACPTAAGACGCTTGCGACAGCAAAAGAACTACCA 1001
211 euArGLuTYrGLInseryVallys..ArgInIllePheserAspyr 226
|||||: : : : :
1000 TCAGGGTGAAAGGCAACCTCAAAAATGGAGAAAAATTTTGGCAACCTAC 951
227 GlnSerGLuValAspIleTYrAsnArqIleaaZspGLueu 240
|||||: : : : :
950 TCATCTGACAAAGGCTAATATCCAGATCTCATGTAAGACT 909

seq_name: /cgn2_6/prodata/1/pna/us10_NEW_COMB.seq.us-10-109-551-1
seq_documentation_block:
Sequence 1, Application us/10109551
GENERAL INFORMATION:
APPLICANT: DUNNE, PATRICK W
TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
FILE REFERENCE: TMK.207US
CURRENT APPLICATION NUMBER: US/10/109,551
CURRENT FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: 60/280,549
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 78056
TYPE: DNA
ORGANISM: Bos taurus
US-10-109-551-1
```



```

104 TYRSEPRONHS...PTQYTGSLGSLGLVALSERALALEUGLYGTYL 119
123 TATTCACCTTCAATTTCTTATGAAATCATCAGACTCATACATATAAAGT 172
119 EROTYGSEGINLLETYGTYGTYPTFYARVALASNHEGTYLLEA 136
173 ACCAACCACAGTAGCAAAAGT.....GACACTTCAAGTAGTACCT 213
136 SPGLUARGLEUHLISARGASNAARGLUYF.....ARG 146
214 CAGAG.....TCGGAAGAAGAAATTAAGTCGTCTATTTGGCCAAA 254
147 ASPARGYTYTYARGASNLLEUASNLLEALPROALAGLUASRGTYTAR 163
255 AATGATTTACTATTAAGACCATCAATT.....TCAGAGGATTAACATAA 298
163 GLEUHLAIGLYRHERPOROASPHISGLNALITPRAGLUGLUPOTRI 180
299 GAAATATCATTTGTTCGCGAACAAGAAATCATCTGCCAAAGAATGACT 348
180 LEHLHIALAPROGLNGLCYSGIYASNSERSERATGTHRIETHRGY 196
349 ACTAADAAGCA.....TTATTTGGAA 374
197 ASPTHYASNLGLUTHRGINSNLSEUERTHRIETYLEUARGI 213
375 GATACACTCAAGAGAGACTCA.....TATGTCAGAA 406
213 UTYGINSERLYSVALYSARGINLEPHESETPRYGINSERGIUV 230
407 GGTACCCCTCGAAGGCTAAACGAAATATTAAGAGATCAATTTTTCAAAT 456
230 ALASPILETYRANARG 235
457 TTGACTACTTCAAGAG 473

seq_name: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq:US-10-109-551-9

seq_documentation_block:
? Sequence 9, Application US/10109551
? GENERAL INFORMATION:
? APPLICANT: DUNNE, PATRICK W.
? APPLICANT: PIEDRAHITA, JORGE
? TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
? FILE REFERENCE: TAMK-207US
? CURRENT APPLICATION NUMBER: US/10/109,551
? PRIOR FILING DATE: 2002-03-28
? PRIOR APPLICATION NUMBER: 60/280,549
? PRIOR FILING DATE: 2001-03-30
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 9
? LENGTH: 830
? TYPE: DNA
? ORGANISM: Cervus elaphus
US-10-109-551-9

alignment_scores:
Quality: 80.00 Length: 124
Ratio: 1.194 Gaps: 6
Percent Similarity: 54.032 Percent Identity: 24.194

alignment_block:
US-09-528-682-1 x US-10-109-551-9 ..
Align seg 1/1 to: US-10-109-551-9 from: 1 to: 830

114 Serialaleuglyglyllleprotyrserglnletyrglytptyrargva 130
:::|||||:::|||||:::|||||:::|||||:::
397 GAGGAGGAGGAGGAGGCTC.....GGTGGCTACTTCT 428

```



```

130  LasnPhcglyValIleAspGluArgLeuHisArgAsnArgIuTyArgA 147
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
429  GGGAGTGCATGAGCAGCGCTTATCATTTGGCAATGACATATGAGG 478
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147  spArgTyTyArg...AsnLeuAsnIleAlaProAlaGluAspGlyTy 162
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
479  ACCGTACTATCCGTAACATGACCTTACCCCAACCAAGTGACTAC 528
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163  ArgLeuAlaGlyPheProAspHisGlnAlaIlePargGluGluPro 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
529  AGG.....CCAGTGCATCAGTATATATACCAAGAACCTT 563
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
179  PLeHisHisAlaProGlnGlyCysGlyAsn.....SerSerArgT 193
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
564  TGTGCAT.....GACTGTGCATCATCATCAGTCAAGCAACACA 601
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
193  hrlleThGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIle 209
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
602  CAGTCACCAACCAACCAAGGGGAG.....AACTTCACCGAAACT 642
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
210  TyrLeuArgGluTyArgInserTyValIysArgGlnIlePheSerAsp 226
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
643  GACATCAAGATATGAGCGCGTGTGAGCAAAATGTGCATCACCCAGTA 692
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
226  rGlnSerGluValAspIleTy 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
693  CCAGAGACATCCGAGCGCTTAT 714

```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-096-080-25

```

seq_documentation_block:
: Sequence 25, Application US/10096080
: GENERAL INFORMATION:
: APPLICANT: Chesebro, Bruce W
: APPLICANT: Caughey, Byron W
: APPLICANT: Chabry, Joelle
: APPLICANT: Piola, Susette
: TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
: FILE REFERENCE: 50121
: CURRENT APPLICATION NUMBER: US/10/096,080
: CURRENT FILING DATE: 2002-03-11
: PRIOR APPLICATION NUMBER: US/09/823,494
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: 09/128,450
: PRIOR FILING DATE: 1998-08-03
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 25
: LENGTH: 1000
: TYPE: DNA
: ORGANISM: Hamster sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (60)..(824)
: US-10-096-080-25

```

alignment_scores:

Quality:	80.00	Length:	124
Ratio:	1.176	Gaps:	6
Percent Similarity:	54.839	Percent Identity:	24.194

alignment_block:

US-09-528-682-1 x US-10-096-080-25 ..

Align seg 1/1 to: US-10-096-080-25 from: 1 to: 1000

```

114 SerAlaLeuGlyGlyIleProTySerGlnIleTyArgIleTyTyArgVa 130
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
417  GCCGTGGGGGGGCGCTT.....GCTGCTACATGCT 448

```

```

130  LasnPhcglyValIleAspGluArgLeuHisArgAsnArgIuTyArgA 147
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
449  GGGAGTGCATGAGCAGCGCCATGATGATTTGGCAATGACTGGGAGG 498
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
147  spArgTyTyArg...AsnLeuAsnIleAlaProAlaGluAspGlyTy 162
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
499  ACCGTACTATCCGTAACATGAAACCGTACCTTACCAAGTGATATAC 548
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163  ArgLeuAlaGlyPheProAspHisGlnAlaIlePargGluGluPro 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
549  CGG.....CCAGTGCATCAGTATATATACCAAGAACACTT 583
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
179  PLeHisHisAlaProGlnGlyCysGlyAsn.....SerSerArgT 193
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
584  TGTGCAT.....GATTGTCAACATCATCACCATCAAGACACACA 621
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
193  hrlleThGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIle 209
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
622  CAGTCACCAACCAACCAAGGGGAG.....AACTTCACCGAGACC 662
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
210  TyrLeuArgGluTyArgInserTyValIysArgGlnIlePheSerAsp 226
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
663  GACATCAAGATATGAGCGCGTGTGAGCAAAATGTGTACACCCAGTA 712
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
226  rGlnSerGluValAspIleTy 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
713  TCAGAGAGATCCCGAGCGCTTAC 734

```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-144-781-41

```

seq_documentation_block:
: Sequence 41, Application US/10144781
: GENERAL INFORMATION:
: APPLICANT: SUBRAMANIAN, Gangadharan et al.
: TITLE OF INVENTION: MOUSE ORTHOLOGS OF HUMAN DRUG TARGET
: TITLE OF INVENTION: GENES, PROTEINS ENCODED BY THESE MOUSE GENES, AND USES
: FILE REFERENCE: C1001236
: CURRENT APPLICATION NUMBER: US/10/144,781
: CURRENT FILING DATE: 2002-05-15
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 41
: LENGTH: 2701
: TYPE: DNA
: ORGANISM: Mus musculus
: US-10-144-781-41

```

alignment_scores:

Quality:	80.00	Length:	209
Ratio:	0.792 <td>Gaps:</td> <td>9 </td>	Gaps:	9
Percent Similarity:	48.325	Percent Identity:	20.574

alignment_block:

US-09-528-682-1 x US-10-144-781-41/rev ..

Align seg 1/1 to reverse of: US-10-144-781-41 from: 1 to: 2701

```

12  ProProAspGluIleLysArgSerGlyGly.....LeuMetProAr 25
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2380  CCTACTGATGAACCTGACAGGTTGAGTGCAGCGCGAATTTTGGGCA 2331
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25  gGlyHisAsnGluTyArgPheAspArg.....GlyThrGlnM 37
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2330  GCCCTATGCCGTAGCCCTGATGCGCAAAAGGTTTCCCGACGAGGAGC 2281
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
37  eAsnIleAsnLeuTyArgHisAla..... 45
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2280  TTGCATCCCGCATCTATGACACCTCATATCCAGTATGATTTATTCAT 2231
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
46  ..ArgGlyThrGlnThrGlyPheValArgTyArgAspAspIleTyValSe 61
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```

1605 ACTCGGCTCTCTCAGGTAACGAGTATTCATATATAAAGAAACTCTACA 1654
      38 snlleasneutyrasphisaalargcluythrglnthgclpheyvalarg 54
      :: |||::|||:: |||
1655 CGACCAACATCTACTCGGACTAC.....TTTAACAG 1686
      55 tyraspargcluytyvalserthserleusertleuargseralnhisale 71
      ::|||::|||:: |||::|||::|||::|||::|||::|||::|||::|||
1687 TTTGACAGCGGACACTATTCCAAAGTTATGCGCGTTTCAACACAGTAGCT 1736
      71 ualagclnserileusergcluyserthtyttrile..... 85
      | .....TATAAACCCGTTATGCGGACTC 1762
1737 T.....TATAAACCCGTTATGCGGACTC..... 86
      86 .....tyvallealathralproasmetpheasnyalasn 98
      :: |||::|||::|||::|||::|||::|||::|||::|||::|||
1763 AAAAGAGCACACTGATTTCAAGTCGGGCTTACATC..... 1800
      99 aspvallleuglyaltyrserprohisarotryglunglvalserl 115
      1801 .....AACACAGCACCGCAAGACCCGAAAC 1826
      115 aleuglyglileprotyrserglniletyrglyttrtyr...Argvala 131
      | ||||| |||::|||::|||::|||::|||::|||::|||::|||::|||
1827 ATGGGAGGAGTATGCCCAACATGCCGAGCGGTTCTTCGGACAGTATCA 1876
      131 snpheglyvalleaspgluargleuhisargasnarugcluytyrargsp 147
      ||:: |||::|||::|||::|||::|||::|||::|||::|||::|||
1877 ACAC.....AATATATAAGAC 1893
      148 argtytytyrargasnleuasnillealproalagluaspglytyrargle 164
      ||||| |||::|||::|||::|||::|||::|||::|||::|||::|||
1894 CGATACATGCTGGACACTAAC...GGACGTTACGCGGTCATCATCAGTPT 1940
      164 ualagly.....pheproproasphinsglnalatrpa 175
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1941 TATAGTGACAAACGCTGGGATTCCTCCCTTCATCTCTCGCCGGATGGA 1990
      175 rg.....glugluptrotprllehisnlsalaproglnglycysglnasn 189
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||
1991 ACCTTGACACAGAACCATCTCTTAAGAANAACGCCGAAANAATGCAATAC 2040
      190 serSer.....Argthrllethgclyaspthrcysasnglcluth 203
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||
2041 GGTACACTTGAAACTCCGCGCTTCTGGGACAAATGAGTAATACAGATAC 2090
      203 rglnasn 205
      |:::|:::|
2091 CAAAGAC 2097

seq_name: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-09921-59
seq_documentation_block:
? Sequence 59, Application PC/TUS0209921
? GENERAL INFORMATION:
? APPLICANT: INCYTE GENOMICS, INC.
? APPLICANT: DAFEO, Abel
? APPLICANT: JONES, Anissa L.
? APPLICANT: TRAN, Alanna-Phung B.
? APPLICANT: DAHL, Christopher R.
? APPLICANT: GIERTZEN, Darryl
? APPLICANT: CHINN, Joyce
? APPLICANT: DUFOUR, Gerard E.
? APPLICANT: HITLMAN, Jennifer L.
? APPLICANT: YU, Jimmy Y.
? APPLICANT: TUASON, Olivia
? APPLICANT: YAP, Pierre E.
? APPLICANT: AMSHEY, Stefan R.
? APPLICANT: DAUGHERTY, Sean C.
? APPLICANT: DAM, Tam C.
? APPLICANT: LIU, Tommy F.
? APPLICANT: NGUYEN, Duy-Viet An
? APPLICANT: KLEBEFELD, Yael

```



```

167 heProboAphisglnAlaIArgglnuProTyr..... 179
    ::      ::|||::|||::|||::|||::|||::|||
420 ATCAAGACAGGAGCAATCTGCGCGGACGACAGTATGAAGAAACGAC 469
180 .....IleHisHisAlaProGlnGlyCysGlyAsnSerSer 192
    ::|||      ::|||      ::|||      ::|||
470 CGGCGGAAGTCTGCTCATCCCATCCATCCAGGCGCATCACAGGTCAGCG 519
192 gThrIleThrGlyAspThrCysAsnGlnGluThrGlnAsnLeuSerThr 209
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
520 GTCTATCATGTGTGATTTAGCA.....GAAGCAAGAGCGCTGTCTC 563
209 leTyrLeuArgLysThrGlnSerLysValLysArg 220
    ::|||      ::|||      ::|||      ::|||
564 TCTGTCTTTCAGTTTCCAGTTTCCAGTTTCTTATTCAGAAA 598

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-26953

seq_documentation_block:
; Sequence 26953, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; .APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26953
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Neurospora crassa
US-60-360-039-26953

alignment_scores:
      Quality: 79.00      Length: 267
      Ratio: 0.731      Gaps: 13
Percent Similarity: 40.449      Percent Identity: 20.599

alignment_block:
US-09-528-682-1 x US-60-360-039-26953 ..

Align seg 1/1 to: US-60-360-039-26953 from: 1 to: 769

1 AsnGlyAspArgLeuTyrArg.....AlaAs 9
::|||::|::|::|::|::|::|::|::|::|
15 CACGGGCAAGGACATTACCATCATTTGCGGCGTCCAAAGCCGCTATGCGG 64
9 pSerArgProAspArgGlnIleLysArgSerGlyGlyLeuMetProArg 26
|||      |||::|::|::|::|::|::|::|::|::|
65 GAGCGACCATCTCTGAAGCATCCGACGACAGACAAGAGAGTCCGTGAACAAG 114
26 LysHisAsnGlu..... 29
::|||::|::|::|::|::|::|::|::|::|
115 CACCACTGATGTGCATTCTCCGGTGAAAGCTGCGACACAGGTACGAGC 164
29 ..... 29

165 GATATCGATGATGCGCCTTTTGACCAACAACGATTACCAACTGTTT 214
30 .....TyrPheAspArgGlyThrGlnMetAsnIleAsnL 41
|||      ::|::|::|::|::|::|::|::|::|
215 GTTCCCATCGACAGTGAAGTGTGCGGAGTACACCAAGCAAGCCAGC 264
41 euTyrAspHisAlaIArgLysThrGln.....ThrGly 51
|||::|::|::|::|::|::|::|::|::|
265 TCTATCATGTGCGCAAGCGACGCGACTTTCCTCCCTTGTCCCTTGGCCAA 314

```



```

52 PheValArgTyrAspArgGlyTyrValSerThrSerLeuSerLeuArgSer 68
   |||||
315 TTTGTCCGC.....GGCGAGCTTGCGAGAGCTTGCGCTCCGCCAA 355
68 rAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrTyrI 85
   :
356 ACCC..... 359
85 leryValIleAlaThrAlaProAsnMetPheAsnValAsnAspValLeu 101
   :|||
360 .....TACACGTAAACCTCTCTC 380
102 GlyValTyrSerProHisProTyrGlnGlnGlu..... 112
   |||
381 GGAGGCTGCACCCCTATCACCCACAGCCAGCTCTACTGCTGCAGTA 430
113 .ValSerAlaLeuGlyGlyIleProTyr...SerGlnIleTyrGlyTyrPT 128
   :|||
431 CTTGGGCTGTTGGCGAAGTCCCTATGCGGCCAGCGCTACGACAGT 480
128 yArgValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGlu 144
   |||
481 ATGTGATCTCGACCGATATCTGTACCGCTTGTGATGCGTAGAGATCAT 530
145 .TyrArg....AspArgTyrTyrArgAsnLeuAsnIleAlaProAlaG 159
   |||||
531 GTACAGATCAGTACCGGTTCTTCTCTATCAGATACTACTGCGCTCG 580
159 lAsp.....GlyTyrArg 163
   :|||
581 ATTCTGCACAGCACCACCCAGACATCACCTCCACAGGCTATCA 630
164 LeuAlaGlyPheProProAspHisGlnAlaTPrArgGluGluProTPrII 180
   :|||
631 GCTTCT....CAACCTGTGCACCGAGAGCTCAAGAGAG.....GT 668
180 eHisHisAlaProGlnGly.....CysGlyAsnSerSerArg 192
   :|||
669 TACCATAGACTTCAGGCGATGACGCTCAAGGCGTACAGAGAGCGG 717
seq_name: /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-109-551-7
seq_documentation_block:
; Sequence 7, Application US/10109551
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; TITLE OF INVENTION: SPONGIFORM ENCEPHALOPATHIES
; FILE REFERENCE: TAKK:20705
; CURRENT APPLICATION NUMBER: US/10/109,551
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Odocolleus hemionus hemionus
US-10-109-551-7

```

```

alignment_scores:
  Quality: 79.00      Length: 124
  Ratio: 1.179      Gaps: 6
  Percent Similarity: 54.032      Percent Identity: 24.194

```

```

alignment_block:
US-09-528-682-1 x US-10-109-551-7
Align seg 1/1 to: US-10-109-551-7 from: 1 to: 771

```

```

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrArgVal 130
   |||||
367 GCAGTGTAGGGGCTC.....GGTGGCTACATGT 398
130 lAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGlyTyrArg 147
   :|||
399 GGGAAGTCCATGAGCAGGCTCTTATACATTTTGGCAATGACATGAGG 448
147 sparGlyTyrArg...AsnLeuAsnIleAlaProAlaGluAspGlyTyr 162
   |||||
449 ACCGTACTACTCGAAGAACATGTACCGTTACCCCAACCAAGTGTACTAC 498
163 ArgLeuAlaGlyPheProProAspHisGlnAlaTPrArgGluGluProT 179
   |||
499 AGC.....CCAGTGCATCAGATATATACCGAAGACCTT 533
179 rIleHisHisAlaProGlnGlyCysGlyAsn.....SerSerArgT 193
   :|||
534 TGTGCAT.....GACTGTGCACATTCACAGTCAAGACAGACA 571
193 hIleThrGlyAspThrCysAsnGluGluTyrGlnAsnLeuSerThrIle 209
   |||||
572 CAGTACACACACACACACAGGAG.....AACTTCACCGAACT 612
210 TyIleuArgGlyTyrGlnSerIleValIleArgGlnIlePheSerArgT 226
   :|||
613 GACATCAAGATGATGAGCGAGCTTGTGAGCAATGTGCATACCCAGTA 662
226 rGlnSerGluValAspIleTyr 233
   |||
663 CCAGAGAGAAATCCAGGCTTAT 684
seq_name: /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-109-551-5
seq_documentation_block:
; Sequence 5, Application US/10109551
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; TITLE OF INVENTION: SPONGIFORM ENCEPHALOPATHIES
; FILE REFERENCE: TAKK:20705
; CURRENT APPLICATION NUMBER: US/10/109,551
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 830
; TYPE: DNA
; ORGANISM: Odocolleus virginianus
US-10-109-551-5

```

```

alignment_scores:
  Quality: 79.00      Length: 124
  Ratio: 1.179      Gaps: 6
  Percent Similarity: 54.032      Percent Identity: 24.194

```

```

alignment_block:
US-09-528-682-1 x US-10-109-551-5
Align seg 1/1 to: US-10-109-551-5 from: 1 to: 830

```

```

114 SerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTyrArgVal 130
   :|||
397 GCAGTGTAGGGGCTT.....GGTGGCTACATGT 428
130 lAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGlyTyrArg 147
   :|||
429 GGGAAGTCCATGAGCAGACCTCTTATACATTTTGGCAATGACATGAGG 478

```



```

147 spargTYTYTArq...AsnleuasnlealaproalaglualaspolyTyr 162
|||||
479 ACCGTACTATCGTAACATGATACCGTACCCACCAACGATGACTAC 528
163 ArgleualaglypneProAspHisGlnalatrparglugluProTr 179
529 AGG.....CCAGTGATCAGTATATAACCAACACACCTT 563
179 PLEHSHISAlaproglnglyCysGlyAsp.....SerSerArgT 193
564 TGTGCAT.....GACTGTCAACATTACACTCAAGCAACCA 601
193 hrlethrGlyAspThrcysasnugluThrGlnAsnleuSerThrIle 209
602 CAGTCACCAACCAACCAAGGGGAG.....AACTCACCAGAACT 642
210 TyrlleuargluTyrGlnSerLysValLysArgGlnIlePheSerAspTyr 226
643 GACATTAAGATGATGACGAGTGTGTGAGCAATGTGCATACCCACTA 692
226 rGlnSerGluValAspIleTyr 233
693 CCAGAGAGAAATCCAGGCTTAT 714
seq_name: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq:us-10-106-698-580

```

```

seq documentation block:
; Sequence 580, Application US/10106698
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentln Ver. 3.0
; SEQ ID NO 580
; LENGTH: 982
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-580

```

```

alignment_scores:
  Quality: 79.00      Length: 178
  Ratio: 1.013      Gaps: 6
  Percent Similarity: 43.820      Percent Identity: 19.663

```

alignment_block:

US-09-528-682-1 x US-10-106-698-580/rev ..

Align seg 1/1 to reverse of: US-10-106-698-580 from: 1 to: 982

```

83 TyrTyrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAs 99
|||||
955 TTTTGTGTTTTTTTTTGTGCTGGAACCAACATTTATGTGACACCTCT 906
99 pValleuglyValTyrSerProHis..... 107
905 GCTCCTCGGAATCTATTCAGTAGATCTTTGCCGAGGCAATCACAATTC 856
108 .....ProTyrGluGlnGluValSerAlaLeu 116
855 ACACAGGGGTGTGACACGACTCCCTCGCTACCTGCGCTGTGGCAGGG 806
117 GlyGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhgI 133

```

```

805 GCAGAGAGTGAATGGCTCT.....TCCTGGGCAAGGCCACAGTTAA 765
133 YValIleAspGluArgLeuHisArgAsnArgGluTyrArg..... 146
764 CCGGTGACAAATGCGAGCCATAGGAATTCATCTTTAGAAAAAACGAA 715
147 .....AsparGlyTyrTyrArgAsnleuasnlealaproalaglualsp 160
714 ACTGATATATAAAAAAACAACCAAACTGAGTATATAAACTCAGAGTGT 665
161 GlyTyrArgleualaglypne..... 167
664 TCCAGCACTATCTCGGGGTTTAAATATAAAGGCATCATGAGAAAAAC 615
168 .....ProProAspHisGlnalatrP. 174
614 TTAATAAGATTAACAGCAGCAGCCGCCACCTCCAGGCACTGTCATGG 565
175 .....ArgGluGluProTyrPheHisHisAlaProglngly 186
564 GGTGGGGCCAGCCCCGCTGAGCGCGGGGCTCCTCCTTGAGCAGCAGG 515
187 Cys.....GlyAsnSerSerArgThrI 194
514 TGCAGGCCCGCCGCTGAGCAGCGGAGCAGCAGCGCCAGCAGCAGCAG 465
194 eThrGlyAspThrcysasnugluThrGlnAsn 205
464 CGCCAGCAGCAGCTGTTCGAGGAGCACCAGGAC 431
seq_name: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq:us-10-105-299-12700

```

```

seq documentation block:
; Sequence 12700, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 12700
; LENGTH: 6078
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12700

```

```

alignment_scores:
  Quality: 79.00      Length: 285
  Ratio: 0.530      Gaps: 15
  Percent Similarity: 52.281      Percent Identity: 21.053

```

alignment_block:

US-09-528-682-1 x US-10-105-299-12700/rev ..

Align seg 1/1 to reverse of: US-10-105-299-12700 from: 1 to: 6078

```

7 ArgAlaSpSerArgProProAspGluIleLysArgSerGlyGlyLeuMe 23
|||||
1723 CAAGCAATATAAAGAGATACAAACAATGAGAGAACATTCATCGCAT 1674
23 tProArgGlyHisasnugluTyrPheAspArgGly.....ThrGlnMeta 38
1673 GGGTAGG...AGCAATCAATATCGTGAATAAGCCATACGCCCCAGAGTA 1627
38 snleuasnleuTyrAspHisAlaArgGlyThrGlnThrGlyPheVal... 53
1626 ATTATACATTCATCAATGTCATCCCATCAACGTCATCAACATGCTTCTTCAC 1577

```



```

54 ArgTyrAspAspGlyTyrValSerThrSerLeuSer...LeuArgSerAl 69
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1576 AGAATTGGAAATAACTTTAAAGTTTCATATGGAACCAAAAAAGAGCCC 1577
69 ah1sLeuAlaGly..... 73
1526 GCATCCGCCAAGCAATCCTAAGCCAAAGAAACAAGAGCTGGAGCATCACA 1477
74 .....GlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyr 86
1476 CTACCTGACTTCAAACTATCTACAGGCTACAGTAACCAAAACAGCATG 1427
87 ValIleAlaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyVal 103
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1426 GTTACTGTTACCAAAACAGAGATATAGTCAATGGAACAGACAGAGCCCT 1377
103 TyrSerProHisProTyrGln.....G 112
1376 CAGAAATATAGCCGATATCTACAACTATGATCTTTGACAAACCTGAG 1327
112 LuValSerAlaLeuGly...GlyIlePro.TyrSerGlnIleTyrGlyTyr 127
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1326 AAAAACCAAGCAATGGGGAGAGATCCCTGTTAATATA.....TG 1286
127 P TyrArgValAsnPheGlyValIleAspArgLeuHisArgAsn.... 142
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1285 GTGCTGGGAAACTGGCTAGCCATATGTACAAAGCTGAAACTGGATCCCT 1236
143 .....ArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
1235 TCCTTACACCTTATACAAATAATCATCAAGATGATTAAGACTTTAAAC 1186
155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAsn 171
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1185 GTTAGACCTTAA.....AACCA 1170
171 sgnAlaTrpArgGluGluProTrpIleHisAlaProGln.GlyCys 187
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1169 TAAAAACCCTAGAAGAAACCTTAGCCCTTACCATTCAGGACATAGGCATG 1120
188 GlyAsnSer.....SerArgThrIleThrGlyAspThrCysAsnGlu 202
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1119 GCGAAGGACTTCATGTCTAAACACCAAAAGCAATGGCACAAGAACAA 1070
202 uThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer..... 216
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1069 AATTGACAATTGGGATCTAATTAACCTAAAGAGCTCTGCACAGCAAAAG 1020
217 .....LysValIysArgGln.....IlePhe 223
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1019 AAACCTACATCAGATGACAGGCAACCTACAAATGGGAAAAATTTC 970
224 SerAspTyrGlnSerGluValAspIleTyrAsnArgIleArgAspGluLe 240
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
969 ACAACCTACTCATCTGACAAAGGCTAATATCCAGATTCATCAATGAACT 920
240 u 240
919 C 919
seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-41301
seq_documentation_block:
; Sequence 41301, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A

```

```

; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41301
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Xylella fastidiosa
; US-60-360-039-41301

alignment_scores:
  Quality: 78.50      Length: 210
  Ratio: 0.853      Gaps: 10
  Percent Similarity: 43.810      Percent Identity: 25.714

alignment_block:
US-09-528-682-1 x US-60-360-039-41301 ..
Align seg 1/1 to: US-60-360-039-41301 from: 1 to: 1050

4 ArgLeuTyrArgAlaAsp.....SerArgPr 12
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
381 CGGCTACTCCACAGATCGTGTGCATGAGCGCTACGCTGCGCATCC 430
12 opProAspGluIleLys.....ArgSerGlyLeuMetProArgGlyH 27
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
431 GCCACGACGAAATCAACACTGCGCGCTGCGCGCGCTGTGCGCGGCC 480
27 IsAsnGluTyrPhe..... 31
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
481 ATCAACACCTACCTGCGCGCTGCGCACCTGGAAAGCGCGCGGCAAGAA 530
32 ...AspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaArgI 47
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
531 GGTGCGCGTGTGCGCATCGCGCGCTGCGGCACATGCGCATCAAGCTGG 580
47 yThrGlnThrGlyPheValArgTyrAspAspGlyTyrVal.SerThrSer 63
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
581 CGCACGCGAATGGCGCGCGACGT.....GTTGCGTTCACCACTCG 621
64 LeuSerLeuArgSerAlaHisLeuAla..... 72
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
622 GAATCCAAAGGCGAGCGCGCATGCGCTGGGTCCGACAGAGTGATGT 671
73 .....GlyGlnSerIleLeuSerGlyTyrSerThrTyrTyrI 85
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
672 CTCGCGCAATCTCGGATGCGCGCGCATGCGCGCACAGAGCTTCGACTCA 721
85 LeuValIleAlaThrAlaProAsnMetPheAsn..... 96
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
722 TCTGAAACAGGTGGCCGCGCGACGACCTCAAGCGCTTCATGAGCGCTG 771
97 .....ValAsnAspValLeuGlyValTyrSerProH 107
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
772 CTCAGGCGGACAGCGGACAGATGAGCGCGGTGCGCGCGCGGTGCGGCA 821
107 sProTyrGlnGluVal..... 113
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
822 CCCC...TCGCGGAACTTTCACACTTCATCTCAAGCGCGGACGCTGG 868
114 .....SerAlaLeuGlyIleProTyrSerGlnIleTyrGlyTyrTyr 128
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
869 CCGGTTCGCAATGCGCGCATCCCGACAGCGAGAGATGCTGACTTC 918
129 ArgValAsnPheGlyValIle...AspGluArgLeuHisArgAsnArgI 144
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
919 TGGCGGAGACAAACATCGTGGCGGACATGAGCTGATCCGCGAGAGGA 968
144 uTyrArgAspArgTyrTyrArgAsnLeu 153
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
969 AATCAACGACGCGCTATGAGCGCATGCTC 996
seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-24895

```



```

seq_documentation_block:
; Sequence 24895, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24895
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Methanobacterium thermoautotrophicum
US-60-360-039-24895

```

```

alignment_scores:
Quality: 78.50 Length: 175
Ratio: 1.061 Gaps: 11
Percent Similarity: 42.286 Percent Identity: 24.000

```

alignment_block:

US-09-528-682-1 x US-60-360-039-24895/rev ..

Align seg 1/1 to reverse of: US-60-360-039-24895 from: 1 to: 1272

```

85 ILETVALLIALLAATHRALAProasmerPheasnValasprValle 101
:::|||||::: ||| |||:::|||||::: ||
585 CTCATGCGAAGTGGATCCCT.....TTCAAGCATATCCGACCTCT 542
101 ugiValTySer.....ProHisProt 109
|||||::: ||| |||:::|||||::: ||
541 GGGCTCCCGCAACATCATCAATTCATGCCCTGAGCCATCTCTG 492
109 YRGUGINGUVALSERALALEUCLYGLIYLEPROTYRSEGINILETYR 125
:::|||||::: ||| |||:::|||||::: ||
491 GTCAGTAT.....GTAAATGGCGGTACCGTTTATATTCATATAT 451
126 GLYTPRYARVALASNPHE...GLYVALLIASPGLUARGLUHSAR 141
|||||::: ||| |||:::|||||::: ||
450 GGATTCATTCCTGTGGCTGCGAGGAGCTCGTGCAG..... 415
141 GASNARGGLUTYRARGSPARGTYRTRYRARGASNLEUASNILEALPROA 158
|||||::: ||| |||:::|||||::: ||
414 .....GGCC 411
158 IAGLUASPGLYTYRARGLEUALGLYPHEPROPROASPHISGLIALTPR 174
|||||::: ||| |||:::|||||::: ||
410 CTGAGAGCGGGGTATAC.....CACACACACACTGG 380
175 .....ArgGLUProTPRILEHis_His.....A 183
|||||::: ||| |||:::|||||::: ||
379 CCTCAATGCAACCCTCAGCGCATGCTCCGAGCATCATATATGTCA 330
183 LAPROGLINGLYCYSGLYASNSESERARGTHRIETHRGLYASPTHCYS 199
|||||::: ||| |||:::|||||::: ||
329 TCCGAGTGCCTTGC.....AAGGAGTGC 307
200 ASNGIUGLUHARGINASNLEUSETHRIILETYRLEUARGGLUTYRGINSE 216
|||||::: ||| |||:::|||||::: ||
306 CTGTGGCTGTTACACATGTTTACCATTTTCTAT...GGCCTTCAGGA 260
216 RLYEVALYASRGLINLEPHESEASRYRGLNSERGLUVALASPLIET 233
|||||::: ||| |||:::|||||::: ||
259 TGAAGCTACAGCGCGGTTTCATAGCCACCTATGAGTCTTATGACATGCT 210
233 YRASNARGILEARGASPLULEU 240

```

```

209 ATTCAGGATCTTCGAGATGTC 187

```

seq_name: /gcn2.6/plodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-34717

```

seq_documentation_block:
; Sequence 34717, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 34717
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Ferroplasma acidarmanus
US-60-360-039-34717

```

```

alignment_scores:
Quality: 78.50 Length: 263
Ratio: 0.649 Gaps: 12
Percent Similarity: 46.008 Percent Identity: 22.053

```

alignment_block:

US-09-528-682-1 x US-60-360-039-34717 ..

Align seg 1/1 to: US-60-360-039-34717 from: 1 to: 1299

```

16 ILELYASRGSLYGLYLEUETPROARGLYHISASNGLUYRPHAS 32
|||||::: ||| |||:::|||||::: ||
322 ATAAAGAAATCTGTGGTTCATTTGAAG.....CTAAGATTTCAAA 365
32 PARGGLYTHGLMETASNLEASNLEUTYRASPNSIALARGLYTHG 49
:::|||||::: ||| |||:::|||||::: ||
366 TTCGGGAGTGAAGCCACATGATGCATAGCGTCGCAAGGAGATTC. 414
49 IETHRG.....PHEVALARGTYRASPARGLYTYRVALISERTHR 62
|||||::: ||| |||:::|||||::: ||
415 .ACGGGGAAGAAAGTACATTGTAAAAATGGAAGCGGATAC..... 453
63 SERLEUSERLEUARGSERALANISLEUALAGLYGLNSETILLEU..... 77
|||||::: ||| |||:::|||||::: ||
454 .....CACGGTGCACATGACTATGCCCTGATAA 482
78 .....SERGLYTYRSETHRTYRTRYRILETYRVALILEALITHALAR 92
|||||::: ||| |||:::|||||::: ||
483 ATCCGGCAGTGGCACCATGACATTCGAGAGTTCATCTTCGACGAGTAC 532
92 ROASNETHPEASNVALASNPVALLEUCLYVALTYRSEPROHISPRO 108
|||||::: ||| |||:::|||||::: ||
533 CGGAGGAAGTTTCCAGAACAGTGTGTTGGAGATACATGACGCTGAA 582
109 .....TYRGUGINGUVALSERALALEUCLYGLIYLEPR 120
|||||::: ||| |||:::|||||::: ||
583 AGCATTTGAANAAGCTTTCAAGAAACACAGAAATGAATTCCTCGGTAT 632
120 OLYRSEGLIILETYRGLYTRYRARGVALASNPHEGLYVALILE... 135
|||||::: ||| |||:::|||||::: ||
633 TACAGAACCCGTTCTTGA.....AATATCGGTATATAAATC 670
136 .....ASRGUARGLEUHNISARGASNARGGLUTYRARGSPARGTYR 150
|||||::: ||| |||:::|||||::: ||
671 CGAGAAATGATTTTGTGAACSTCTACGGGAAATTAACSTGAATATGTA 720
151 ARGASNLEUASNILEALPROALAGLUASPLIYTYRARG..... 163

```



```

721 TCCCTGCCTATATTGATGAAAGTAATAACCGGCTTTAGATTCCGATTAA 770
||||| ::: :|||:|||||
164 ..... LeuAlaGlyPheProPaspHisGlnAlaTrp.... 174
:::|||||: ||||| :::
771 GGGATATCAGACGACATTCGAGAAATAAACCTGATCTTACCACATATGGAA 820
175 ..... ArgGluGlu 177
|||||:|||||:
821 AAATTTATCGGGGAGCGCCTCTGTAGCGCATTTGGCGGGGTGAAGAT 870
178 ProTrpIleHisHisAlaProGlnGlyCysGlyAsnSerSerArgThrIle 194
||| |||||:||||| :|||: |||||
871 ATTTATGGCGCATATGTAAGCCCCCTAGGGTATGATATACGAACGGGAACGT 920
194 eThhGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleLeu 211
:||||| :||| :|||: :||| :|||
921 TTTCGGGG..... AATCCTTTGACCATGCTCGACGAATGTCAACATC 961
211 euaTrgGluTrpGlnSerIleValLysAlaGlnIlePheSerAspTrpTrpGln 227
|| ::: :|||:||||| :|||: |||||
962 TTGACCACTTAAAGGCGAAG..... GATTATTCA 990
228 SerGluValAspIleTrpAsnArgIleAlaArgAspGluLeu 240
||| :||| :|||: |||: |||
991 GCCATTACTTAATCTATGCGACCGCTTGGAAGAAAACTT 1029

```

```
seq_name: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-138411
seq_documentation_block:
: Sequence 138411, Application US/10027632
```

```

1 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
2 POLYMORPHISMS IN THE HUMAN GENOME
3 FILE REFERENCE: 108827.129
4 CURRENT APPLICATION NUMBER: US/10/027,632
5 CURRENT FILING DATE: 2002-04-30
6 PRIOR APPLICATION NUMBER: US 60/218,006
7 PRIOR FILING DATE: 2000-07-12
8 PRIOR APPLICATION NUMBER: US 60/198,676
9 PRIOR FILING DATE: 2000-04-20
10 PRIOR APPLICATION NUMBER: US 60/193,483
11 PRIOR FILING DATE: 2000-03-29
12 PRIOR APPLICATION NUMBER: US 60/185,218
13 PRIOR FILING DATE: 2000-02-24
14 PRIOR APPLICATION NUMBER: US 60/167,363
15 PRIOR FILING DATE: 1999-11-23
16 PRIOR APPLICATION NUMBER: US 60/156,358
17 PRIOR FILING DATE: 1999-09-28
18 PRIOR APPLICATION NUMBER: US 60/146,002
19 PRIOR FILING DATE: 1999-08-09
20 NUMBER OF SEQ ID NOS: 325720
21 SOFTWARE: FastSeq for Windows Version 4.0
22 SEQ ID NO 138411
23
24 LENGTH: 806
25
26 TYPE: DNA
27 ORGANISM: Human
28 US-10-027-632-138411

```

[illegible]

alignment_block:
US-09-528-682-1 x US-10-027-632-138411

Align seg 1/1 to: US-10-027-632-138411 from: 1 to: 806

89 AlaThrAlaProAsnMet.....PheAsnValAsnAspValLeuGlyVa 103
|||||::|||::: || :: |||::

```

587 GCCACCGTCGCCACCGCTTTATTTTCTTTAAAAAAGTGGGGAC 636
103 lTySerProHIsProTyrgInGluIvalSerAlaLeuGlyGlyIleP 120
   :   |||   |||   ::||| |||   :: ||| ||| |||
637 AAGTCIGATGATGATTACTGGAAGAAGTATTTTCATATGTTGGAAATTC
   :   |||   |||   |||   |||   |||   |||   |||
120 roTySerGlnIleTyrg GlyTPTyTrAlrGylAsnPhcGlyIvalIleAs 136
   :|||::: ||| ::|||::: |||::: |||::: |||::: |||:::
687 TATTTCACAGCATTCACACGCTGGCACAAGTAAGTTTGGAGTATCA 736
136 pGluArgLeuHIsArgAsnArgGlyTyr 145
   :   ||| ||| ||| ||| ||| ||| ||| ||| |||
737 ATTATGCTGTGATCATTCCTCGAGGTAC 764

```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-138412

```

seq_documentation_block:
  - Sequence 138412, Application US/10027632
  - GENERAL INFORMATION:
    - APPLICANT: Wang, David G.
    - TITLE OF INVENTION: Identification and Mapping of Single Nucleotides
      in the Human Genome
    - TITLE OF INVENTION: Polymorphisms in the Human Genome
    - FILE REFERENCE: 108827.129
    - CURRENT APPLICATION NUMBER: US/10/027,632
    - CURRENT FILING DATE: 2002-04-30
    - PRIOR APPLICATION NUMBER: US 60/218,006
    - PRIOR FILING DATE: 2000-07-12
    - PRIOR APPLICATION NUMBER: US 60/198,676
    - PRIOR FILING DATE: 2000-04-20
    - PRIOR APPLICATION NUMBER: US 60/193,483
    - PRIOR FILING DATE: 2000-03-29
    - PRIOR APPLICATION NUMBER: US 60/185,218
    - PRIOR FILING DATE: 2000-02-24
    - PRIOR APPLICATION NUMBER: US 60/167,363
    - PRIOR FILING DATE: 1999-11-23
    - PRIOR APPLICATION NUMBER: US 60/156,358
    - PRIOR FILING DATE: 1999-09-28
    - PRIOR APPLICATION NUMBER: US 60/146,002
    - PRIOR FILING DATE: 1999-08-09
    - NUMBER OF SEQ ID NOS: 325720
    - SOFTWARE: FastSeq for Windows Version 4.0
  - SEQ ID NO 138412
  - LENGTH: 806
  - TYPE: DNA
  - ORGANISM: Human
  - US-10-027-632-138412

```

[illegible]

alignment_block:

Align seg 1/1 to: US-10-027-632-138412 from: 1 to: 806

```

89 AlArthAlaProAsnMet.....PheAsnValAsnAspValLeuIlyA 103
|||||:|||||:|||||:|||||:
587 GCCACGCGGCCAGCGCTTTATTTTCTTTTAAAAAAAAGTGGGGAC 636
|||||:|||||:|||||:|||||:
103 LTyrSerProHisProTyrGluGlnGluValSerAlaLeuGlyIlyLeP 120
|||:|||||:|||||:|||||:
637 AAAGCTGATGATAGTACTTACCTGAAAAGAACTATTTTATATGGTGGAAATTC 666
|||||:|||||:|||||:|||||:
120 roTyrSerGlnIleTyr GLTyrPTyrAcgValAsnPhgIlyValIleAs 136
|||||:|||||:|||||:|||||:
687 TATTTAGAGAGCGCTTACAAGCTGGGCGACCAAGTAAGTATTAGAGATGCCA 736
|||||:|||||:|||||:|||||:
136 pGluArGleuHisArgAsnArgGluTyr 145
|||||:|||||:|||||:|||||:
137 ATTATGGTGGCATCATTCCTCAGGCGTAC 764
|||||:|||||:|||||:|||||:

```



```
seq_name: /cgn2.6/plodata/1/pna/US10_NEM_COMB.seq:US-10-109-551-3
seq_documentation_block:
; Sequence 3, Application US/10109551
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; TITLE OF INVENTION: SPONGIFORM ENCEPHALOPATHIES
; FILE REFERENCE: TANK:207US
; CURRENT APPLICATION NUMBER: US/10/109,551
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 31412
; TYPE: DNA
; ORGANISM: Ovis aries
US-10-109-551-3

alignment_scores:
Quality: 78.00 Length: 124
Ratio: 1.164 Gaps: 6
Percent Similarity: 54.032 Percent Identity: 24.194

alignment_block:
US-09-528-682-1 x US-10-109-551-3 ..

Align seg 1/1 to: US-10-109-551-3 from: 1 to: 31412

114 Serialaleuclgylleprotyrserghilletyrlytrpyrargva 130
      :::::::::::::::::::: ||| |||
22644 GCAGTGGTAGGGGCGCTT.....GGTGGTACAGTCT 22675
      :::::::::::::::::::: ||| |||
130 lasnphg1valli1leaspluarleuhs1argasnarg1utyargva 147
      :::::::::::::::::::: ||| |||
22676 GCGAAGTCCATGAGCAGCGCTGTATACATTTTGGCAATGACATACAG 22725
      :::::::::::::::::::: ||| |||
147 spargtyrtyrarg...Asnleuhsnillealproalagluasrplytyr 162
      :::::::::::::::::::: ||| |||
22726 ACCGTACTACTCGTGAACATGATACCGGTACCCCAAGAGTACTAC 22775
      :::::::::::::::::::: ||| |||
163 Argleuhalaglyrherproarphn1slalatrarglugu1urprotr 179
      :::::::::::::::::::: ||| |||
22776 AGA.....CCAGTGGATCAGTATAGTACCCAGAACAACTT 22810
      :::::::::::::::::::: ||| |||
179 rileh1sh1alaproglnglucysgluasn.....Sersearagt 193
      :::::::::::::::::::: ||| |||
22811 TGTGTCAT.....GACTGTGTCAACATCAGTCAGTCAAGAACACA 22848
      :::::::::::::::::::: ||| |||
193 hrllethrglyasprthrcysasnglugu1uthrglnasnlseurthile 209
      :::::::::::::::::::: ||| |||
22849 CAGTCACACACACACACACAGGGGAG.....AACTTCACCGAAGCT 22889
      :::::::::::::::::::: ||| |||
210 Tytleuargluyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyr 226
      :::::::::::::::::::: ||| |||
22890 GACATCAAGATATGAGCGAGTGGTGGACCAATGTGCATCACCAGTA 22939
      :::::::::::::::::::: ||| |||
226 rglnsergluvalasplletyr 233
      :::::::::::::::::::: ||| |||
22940 CCAGAGAGAAATCCAGGCTTAT 22961
      :::::::::::::::::::: ||| |||

seq_name: /cgn2.6/plodata/1/pna/US09_NEM_COMB.seq:US-09-602-777A-227
seq_documentation_block:
; Sequence 227, Application US/09602777A
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Krieger, Burkhard
```

```
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 227
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(997)
; OTHER INFORMATION: KXA01932
US-09-602-777A-227

alignment_scores:
Quality: 77.50 Length: 77
Ratio: 1.845 Gaps: 2
```



```

714 CACGATAGTATGATGAAGAACTGAAGCCGCTACTGTTAAATCGCGGCTAA 763
77 euserglytyrserthrtyletyrvallealea..... 89
764 TGACAAAGACTCTTGTACGTATGAGGCACTATCTCAAGCGGCTGAT 813
90 .....ThralaProAsnme 94
814 GCTTTCAGCAAGAGGACACTTCTTGGAGAGAAATGCTGATACCA 863
94 tPheAsnValAsnAspValleucllyvaltyrserProHisProtyrGluG 111
864 GTTGGAAATAAATACCGTCTCGGCGCTATGTGATATGACAAATPATA 913
111 lngluValserAlaLeuGlyIleProtyr.....SerGln 123
914 TCTGGACAGCCCTTGACAAATGTAATGCTATGCGCAATAATTCGCTG 963
124 IleTyrlGlyTrp.....TyrArgValAsnPhelGlyValIleAspGluar 138
964 ATTACCATTTTCAGCGCGGTTCGCGCAAGTCGGGATGGTTATGATGT 1013
138 glueHisArgAsnArgGlyIuTyraArgAspArgtyrTyraArgAsnLeuAsn1 135
1014 GCTGGTAACGGAT.....AAGGATGCTTAC.....A 1039
155 leaIleProAlaGluAspGlyTyraArgLeu..... 164
1040 TCCGATCCAAATCAAGACTGTACCGGCTTAAAGATACCCGCTGAGCTTA 1089
165 ..AlaGlyPheProProAsnHisGlnAlaTrpArgGluGluProTrp11 180
1090 GTTCCGGGCGCT.....GAAGAGCAGGCAATGAGCATTTGAGAGTGC... 1131
180 eHisHisAlaProGlnGly.....CysGlyAsnSerSerArgT 193
1132 .....GGCGGCCAGGTGCTTTCGGGTCAATAATAAGTA 1165
193 hr.....IleThrGly 196
1166 CTTTCCAGATTAAAGCATGCAAGCCCGCTTCTTCCGATGTGCGGGG 1215
197 AspThrCysAsnGluGluThrGln 204
1216 GCTATGTGTATCGCGCAGCGCGCAG 1239

seq_name: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:US-60-360-039-41882
seq_documentation_block:
; Sequence 41882, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10652052A
; CURRENT APPLICATION NUMBER: US/60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41882
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Thermoplasma acidophilum
US-60-360-039-41882

```

```

alignment_scores:
Quality: 77.00      Length: 260
Ratio: 0.681       Gaps: 12

```

```

Percent Similarity: 43.462      Percent Identity: 21.538
alignment_block:
US-09-528-682-1 x US-60-360-039-41882
Align seg 1/1 to: US-60-360-039-41882 from: 1 to: 1311.

```

```

12 ProProAspGluIleLysArgSerGlyGlyLeuMetProArgGlyHisAs 28
388 CGCGCGGAGGACACTACTACTCTCC.....CGAACAATGCCCTC 425
28 nGluTyR.....PheAspArgGlyThrGlnMetAsnIleA 40
426 CGAGTACGATCTCAACTGCGCGCTGATGCGGAGATAGCCATAAAGCTTCA 475
40 snLeuTyraPheHisAlaArg..... 46
476 ACACCTTTCACACAGTACAGAAAGATAGGAAAGCCGAGAGGATATTC 525
47 .....GlyThrGlnTh 50
526 TTCAGGATCAACCCGGCTTTGGCATGGCGAATTCGACAGTACCAACAC 575
50 rGlyPheValArgTyraPaspGlyTyR.....ValSerThrSerLeus 65
576 CGGGGGTGCAGAGCAAAAGTTCCGGATAGACCCGGATCCGCCATCTCGG 625
65 erLeuArgSerAlaHisLeuAlaGly.....GlnSer 75
626 CCTACAGAAAGGCCAGGAGAGACGGGATCAGGAGATTCCGTTATACATAG 675
76 IleLeuSerGlyTyriserThrtTyrlTyrlTyrlValIleAlaThrAlaBr 92
676 ATGATCGATGATCGAACAACAGGATCATGTCAAGATGCTGAAGCGTATTC 725
92 OAsnMetPheAsnValAsnAspValleucllyvaltyrserProHisProt 109
726 GAATTTCTTCCGCAATTCAGACAGCATCGGAGAGGAGCGGCGTAGTT 775
109 yrcGlu...GlnGluValserAlaLeuGlyIleProtyrserGlnIle 124
776 TTCATATTCGCGGATGTGGGGGCGCTGGCATTCCTGATGTCCAG... 822
125 TyrglyTyrlTyraArgValAsnPhelGlyValIleAspGluArgLeuHisAr 141
823 ...GGCGAAACGAACTGAGACATTCGGCGACCTGGGTGACGCCGTG... 864
141 gAsnArgGlyTyraArgAspArgTyTyR..... 150
865 ..TTGAAAGAAATTT...GATCGATACCATTTCCGCGATCTGCTCTCGAAC 910
150 ..... 150
911 CGGGACCTATCTGGTTGGCGATGCCGCATCATCTCGGCACAGTCAC 960
151 .....ArgAsnLeuAsn11 155
961 GATGTCACACAGCGCTTCGGGGCAGACAGACATCGGATCACTCAACAT 1010
155 eAlaProAlaGluAspGlyTyraArgLeuAlaGlyPheProProAspHisG 172
1011 AAGGCGCGGCGCTGTACGGCGCAGGACAGATAAATCCCGTGGGAGAGA 1060
172 lAlaLeuTyraArgGluGluProTrpIleHisAlaProGlnGlySerGly 188
1061 GGGTGTAGGCGAGAGAG.....ATCACCGTCACCGGCGATATATGCGAG 1104
189 AsnSerSerArgThrIleThrGlyAspThr 198
1105 AACACGCGACAGG.....ATCGCGCATACG 1128

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-096-080-27

```



```
; Sequence 26535, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26535
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Synecocystis sp.
US-60-360-039-26535
```

```
alignment_scores:
    Quality: 77.00      Length: 201
    Ratio: 0.885      Gaps: 7
    Percent Similarity: 43.284      Percent Identity: 21.891
```

alignment_block:

US-09-528-682-1 x US-60-360-039-26535 ..

Align seg 1/1 to: US-60-360-039-26535 from: 1 to: 2163

```
15 GluIleLysArgSerGlyLeuMetPro.....ArgLysH 27
   ||||| ||| ||||| |||||
1624 GAATTTATCTTCCCGATTTGATCCCGACAGATTGGCGCATCA 1673
   ||||| ||||| ||||| |||||
27 SAANGIUTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspH 44
   ||||| ||||| ||||| |||||
1674 CATGATTAATTTG.....TGGAACTTAACGCCATTTGTCG 1708
   ||||| ||||| ||||| |||||
44 IAlaArgGlyThrGlnThrGlyPheValArgTyrAspArgGlyTyrVal 60
   ||||| ||||| ||||| |||||
1709 GTATTTGGGGGAATTACCCGGTTTACC..... 1737
   ||||| ||||| ||||| |||||
61 SerThrSerLeuSerLeuArgSerAlaHisIleuAlaGlyInSerIleLe 77
   ||||| ||||| ||||| |||||
1738 .....CTCCAACTTAACGAGAACT 1757
   ||||| ||||| ||||| |||||
77 userGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnH 94
   ||||| ||||| ||||| |||||
1758 GAGCGCTATTAATAATTAGCTATTCTACGATGCGGCGATCGCCCTATACC 1807
   ||||| ||||| ||||| |||||
94 eTPheAsnValAsnAspValIleuGlyValTyrSerProHisProTyrGlu 110
   ||||| ||||| ||||| |||||
1808 TAGAGAAATTCGCAACTGTG.....CATTAAGGGGA 1842
   ||||| ||||| ||||| |||||
111 GlnGluValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGlyTr 127
   ||||| ||||| ||||| |||||
1843 CAACACCTAATATACATC..... 1860
   ||||| ||||| ||||| |||||
127 pTyrArgValAsnPheGlyValIleAspGluArgLeu.....HisArgH 142
   ||||| ||||| ||||| |||||
1861 .....ATTTCTTTGGTCAATTTTGGATTTCTGCCATCCGACCTT 1903
   ||||| ||||| ||||| |||||
142 snArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAla 158
   ||||| ||||| ||||| |||||
1904 CCAAGGCTATGCTGCTGCTGTTGAGCCACAGTGGAAATATTTCCCTG 1953
   ||||| ||||| ||||| |||||
159 GlnAspGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTyrPar 175
   ||||| ||||| ||||| |||||
1954 GAGCAGCTTTTACCGCGGAGATGCGGAGCGACGAGATATATATGCG 2003
   ||||| ||||| ||||| |||||
175 gGluGluProTyrPheHisIleAlaProGlnGlyCysGlyAsnSerSera 192
   ||||| ||||| ||||| |||||
2004 G.....GGTAAACACCTTT 2017
```

```
192 rGThrIleThrGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThr 208
   ||||| ||||| ||||| |||||
2018 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2067
   ||||| ||||| ||||| |||||
209 Ile 209
   ||||| ||||| ||||| |||||
2068 ATC 2070
```

seq_name: /cgn2_6/plodata/1/pna/US10_NEW_COMB.seq:US-10-104-047-1307

seq_documentation_block:

```
; Sequence 1307, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1307
; LENGTH: 2405
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1307
```

```
alignment_scores:
    Quality: 77.00      Length: 187
    Ratio: 0.802      Gaps: 11
    Percent Similarity: 51.337      Percent Identity: 23.529
```

alignment_block:

US-09-528-682-1 x US-10-104-047-1307 ..

Align seg 1/1 to: US-10-104-047-1307 from: 1 to: 2405

```
9 AspSerArgProProAspGluIleLysArgSerGlyLeuMetProHr 25
   ||||| ||||| ||||| |||||
1465 GACACGTGCTCCCGCTGACATACATCCCAAGATCTGACCTCCACCCG 1514
   ||||| ||||| ||||| |||||
25 gGlnHis...AsnGluTyrPheAspArgGlyThrGlnMetAsnIleAsnL 41
   ||||| ||||| ||||| |||||
1515 TCACCATGGCACTTACACCATGACCGTGCACGGTCTGCTGCTGCC 1563
   ||||| ||||| ||||| |||||
41 eUTyrAspHis..AlaArgGlyThrGlnThrGlyPheValArgTyrAspAs 57
   ||||| ||||| ||||| |||||
1564 ..TATGATCATCGCGCGTGGCAAT.....GACCA 1590
   ||||| ||||| ||||| |||||
57 pGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisIleuAlaGlyG 74
   ||||| ||||| ||||| |||||
1591 GTGGCCATCAGCTCCAAATTTGAGACCCGGAG.....GATA 1628
   ||||| ||||| ||||| |||||
74 InSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrValIleAlaThr 90
   ||||| ||||| ||||| |||||
1629 TTGCTGTGATCCGAGACTATGGAACTCTCTGTGAGATGTCGCTG 1678
   ||||| ||||| ||||| |||||
91 AlaProAsnMetPheAsnValAsnAspValIleuGlyValTyrSerProH 107
   ||||| ||||| ||||| |||||
1679 GTCCCT.....GATGGCATCGGGCTTCTTACACGAGT 1713
   ||||| ||||| ||||| |||||
107 sProTyrGluGlnGluValSerAlaLeuGlyIleProTyrSerGlnI 124
   ||||| ||||| ||||| |||||
1714 CCAATCATGAGAG.....AGCACCG 1733
   ||||| ||||| ||||| |||||
124 IeTyrGlyTyrTyrArgValAsnPheGlyValIleAspGluArgLeuHis 140
   ||||| ||||| ||||| |||||
1734 TGCCCTCTGCTGATGACGAG.....GGGATCTTT..GAGAACATCCAG 1774
   ||||| ||||| ||||| |||||
141 ArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaBr 157
```


gb_ba:AE0064158	+	8.00	112.30	682.68	11378	AE004158	Vibrio cholerae chcd	gb_htg:LMFLCHR32_22	+	8.00	96.11	5.4e+03	110000	Continuation (23 of 36) o
gb_ba:AE0064158	+	8.00	112.30	682.92	11437	AE006663	Sulfolobus solitarius	gb_htg:LMFLCHR36_24	+	8.00	96.11	5.4e+03	110000	Continuation (25 of 36) o
gb_in:CEH19N07	+	8.00	111.94	714.90	111966	292835	Caenorhabditis elegans	gb_pl:AC006258	+	8.00	96.08	5.5e+03	110469	AC006258 Arabidopsis thal
gb_ba:AE009278	-	8.00	111.59	747.59	12555	AE009278	Agrobacterium tumefaciens	gb_htg:AC001801	+	8.00	96.03	5.5e+03	111187	AP001801 Homo sapiens chr
gb_htg:AC014206	-	8.00	111.49	757.28	12743	AC014206	Agrobacterium tumefaciens	gb_pr:AC026441	+	8.00	95.95	5.6e+03	112532	AC026441 Homo sapiens chr
gb_ba:AE001969	+	8.00	111.35	771.62	13007	AE001969	Deinococcus radiodurans	gb_pr:AC010584	+	8.00	95.90	5.6e+03	113216	AC010584 Homo sapiens chr
gb_ba:AE008334	+	8.00	110.49	861.18	14665	AE008334	Agrobacterium tumefaciens	gb_pr:AC013354	+	8.00	95.86	5.6e+03	113988	AC013354 Homo sapiens chr
gb_ba:AE004755	-	8.00	110.44	866.23	14759	AE004755	Pseudomonas aeruginosa	gb_htg:AC010310	+	8.00	95.73	5.7e+03	116089	AC010310 Rattus norvegicus
gb_htg:AC0109866	+	8.00	109.78	979.73	16884	AC010986	Rattus norvegicus cl	gb_pr:AL15780	+	8.00	95.65	5.8e+03	116878	AL15780 Homo sapiens chr
gb_htg:AC0109866	+	8.00	107.75	1.2e+03	16884	AC010986	Rattus norvegicus cl	gb_htg:AC008457	+	8.00	95.68	5.8e+03	117366	AC008457 Homo sapiens chr
gb_pat:AX076880	+	8.00	107.32	1.3e+03	32870	AX076880	Rattus norvegicus cl	gb_htg:AC009759	+	8.00	95.58	5.8e+03	118444	AC009759 Rattus norvegicus
gb_pr:AC004462	+	8.00	105.04	1.7e+03	31491	AC004462	Sequence 1 from Patg	gb_pl:AC007534	+	8.00	95.55	5.9e+03	118968	AC007534 Genomic sequence
gb_pr:HS366D1	+	8.00	104.08	2.0e+03	36021	AC004462	Homo sapiens Chromos	gb_htg:AF198096	+	8.00	95.45	5.9e+03	120731	AF198096 Homo sapiens chr
gb_htg:AC019927	+	8.00	103.91	2.0e+03	36879	AC019927	Drosophila melanogaster	gb_pr:AC010738	+	8.00	95.40	6.0e+03	121448	AC010738 Homo sapiens BAC
gb_pr:AC005061	-	8.00	103.62	2.1e+03	38413	AC005061	Homo sapiens Chromos	gb_htg:AP004093	+	8.00	95.36	6.0e+03	122218	AP004093 Oryza sativa chr
gb_pl:AC005299	-	8.00	103.23	2.2e+03	40554	AC005299	emericella nidulans	gb_pr:AC016648	+	8.00	95.26	6.1e+03	123946	AC016648 Homo sapiens chr
gb_htg:AC096021	+	8.00	103.03	2.2e+03	41758	AC096021	Rattus norvegicus cl	gb_pr:AC004904	+	8.00	95.14	6.2e+03	126041	AC004904 Homo sapiens PAC
gb_htg:AC110114	+	8.00	102.55	2.4e+03	44615	AC110114	Rattus norvegicus cl	gb_htg:AC099140	+	8.00	95.10	6.3e+03	128511	AC099140 Rattus norvegicus
gb_pl:FC22M8	-	8.00	102.13	2.5e+03	47295	AC020622	Sequence of BAC F22M	gb_htg:AC096871	+	8.00	94.88	6.4e+03	130692	AC096871 Rattus norvegicus
gb_htg:AC006182	-	8.00	101.79	2.6e+03	49636	AC006182	Homo sapiens Chromos	gb_pr:AC079772	+	8.00	94.88	6.4e+03	130743	AC079772 Homo sapiens BAC
gb_pr:AL356515	-	8.00	100.81	3.0e+03	56951	AL356515	Oryza sativa Chromos	gb_pr:AC008971	+	8.00	94.82	6.4e+03	131856	AC008971 Homo sapiens chr
gb_htg:AC017426	-	8.00	100.47	3.1e+03	59704	AC017426	Drosophila melanogaster	gb_htg:AL356692	+	8.00	94.66	6.5e+03	134787	AL356692 Homo sapiens chr
gb_pr:AC000260	-	8.00	100.43	3.1e+03	60023	AC000386	Homo Chromosome 15	gb_htg:AC103239	+	8.00	94.50	6.7e+03	137897	AC103239 Rattus norvegicus
gb_pr:HS1173K17	-	8.00	100.04	3.3e+03	63402	AP000260	Homo sapiens genom	gb_pr:AL157764	+	8.00	94.48	6.7e+03	138251	AL157764 Homo sapiens PAC
gb_htg:AC090157	+	8.00	99.89	3.4e+03	64810	AL050315	Homo sapiens genom	gb_pr:AC009226	+	8.00	94.45	6.7e+03	138778	AC009226 Homo sapiens BAC
gb_htg:AC090157	+	8.00	99.84	3.4e+03	65239	AC090157	Homo sapiens Chromos	gb_pl:AL157821	+	8.00	94.45	6.7e+03	138781	AL157821 Homo sapiens chr
gb_htg:AC100935	-	8.00	99.84	3.4e+03	65239	AC090157	Homo sapiens Chromos	gb_pr:AC009177	+	8.00	94.39	6.8e+03	140064	AC009177 Arabidopsis thal
gb_htg:AC100935	-	8.00	99.71	3.4e+03	66445	AC100935	Mus musculus clone R	gb_pr:AC010253	+	8.00	94.38	6.8e+03	140249	AC010253 Homo sapiens chr
gb_htg:AC110274	+	8.00	99.49	3.5e+03	68482	AC110274	Homo sapiens clone R	gb_htg:AC013376	+	8.00	94.37	6.8e+03	140290	AC013376 Homo sapiens chr
gb_pr:HSU70984	+	8.00	99.46	3.5e+03	68548	U70984	Human Xp22. cosmids U70	gb_pr:AL356107	+	8.00	94.36	6.8e+03	144857	AL356107 Homo sapiens chr
gb_htg:AC090318	-	8.00	99.46	3.5e+03	68753	AC090318	Homo sapiens Chromos	gb_pr:AP000036	+	8.00	94.15	7.0e+03	144857	AP000036 Homo sapiens chr
gb_htg:AC010739	-	8.00	99.38	3.6e+03	69537	AC010739	Homo sapiens Chromos	gb_htg:AC069591	+	8.00	94.12	7.0e+03	145458	AC069591 Homo sapiens chr
gb_htg:AC027823	-	8.00	99.26	3.6e+03	70720	AC027823	Homo sapiens Chromos	gb_pr:AC008766	+	8.00	94.02	7.1e+03	147530	AC008766 Homo sapiens chr
gb_htg:AC011590	+	8.00	99.18	3.7e+03	71582	AC011590	Mus musculus clone R	gb_htg:AC057542	+	8.00	94.02	7.2e+03	148172	AC057542 Fells calus chr
gb_htg:AC025914	+	8.00	99.15	3.7e+03	72450	AC025914	Rattus norvegicus cl	gb_htg:AF166490	+	8.00	93.97	7.2e+03	148355	AF166490 Homo sapiens chr
gb_pl:AC074228	+	8.00	99.05	3.7e+03	72832	AC074228	Mus musculus Chromos	gb_htg:AF162730	+	8.00	93.93	7.2e+03	149354	AF162730 Dario reio chr
gb_pr:AL136159	-	8.00	98.94	3.8e+03	74012	AL136159	Human DNA sequence	gb_htg:AP004193	+	8.00	93.88	7.2e+03	150275	AP004193 Oryza sativa chr
gb_pr:AP001264	-	8.00	98.89	3.8e+03	74550	AP001264	Homo sapiens genom	gb_htg:AL356215	+	8.00	93.87	7.3e+03	150533	AL356215 Homo sapiens chr
gb_htg:AC037461	+	8.00	98.59	4.0e+03	77654	AC037461	Homo sapiens Chromos	gb_htg:AC018889	+	8.00	93.87	7.3e+03	150629	AC018889 Bos taurus clone
gb_pr:AL1590611	-	8.00	98.57	4.0e+03	77888	AL159061	Homo sapiens sequence	gb_pr:AC007719	+	8.00	93.86	7.3e+03	150831	AC007719 Homo sapiens chr
gb_pr:AL139826	-	8.00	98.35	4.1e+03	80419	AL139826	Human DNA sequence	gb_pr:AL354740	+	8.00	93.81	7.3e+03	151828	AL354740 Homo sapiens chr
gb_pr:AC084807	+	8.00	98.23	4.1e+03	81704	AC084807	Arabidopsis thaliana	gb_pr:AL359074	+	8.00	93.79	7.3e+03	152161	AL359074 Homo sapiens chr
gb_pr:AC034726	+	8.00	98.18	4.2e+03	82316	AL354726	Human DNA sequence	gb_htg:AC008688	+	8.00	93.76	7.4e+03	152230	AC008688 Homo sapiens chr
gb_pl:AC007197	-	8.00	98.09	4.2e+03	83363	AC007197	Arabidopsis thaliana	gb_htg:AC013906	+	8.00	93.74	7.4e+03	153386	AC013906 Homo sapiens chr
gb_htg:AC017221	+	8.00	98.05	4.2e+03	83824	AC017221	Drosophila melanogaster	gb_htg:AC087656	+	8.00	93.71	7.4e+03	154067	AC087656 Homo sapiens chr
gb_pr:AP001255	+	8.00	98.04	4.3e+03	83921	AP001255	Homo sapiens Chromos	gb_htg:AC026033	+	8.00	93.68	7.4e+03	154323	AC026033 Homo sapiens chr
gb_htg:AL645826	+	8.00	97.62	4.5e+03	89064	AL645826	Dario reio Chromos	gb_pr:AL355871	+	8.00	93.68	7.4e+03	154588	AL355871 Homo sapiens chr
gb_htg:AC105557	+	8.00	97.55	4.5e+03	89964	AC105557	Rattus norvegicus cl	gb_pr:AC004846	+	8.00	93.68	7.4e+03	154616	AC004846 Homo sapiens PAC
gb_htg:AC109441	-	8.00	97.54	4.5e+03	89990	AC109441	Homo sapiens Chromos	gb_htg:AC026039	+	8.00	93.67	7.4e+03	154630	AC026039 Homo sapiens chr
gb_pr:AC008954	-	8.00	97.53	4.5e+03	90111	AC008954	Homo sapiens Chromos	gb_htg:AC011859	+	8.00	93.67	7.4e+03	154809	AC011859 Homo sapiens chr
gb_pr:AF225900	+	8.00	97.18	4.7e+03	94645	AF225900	Homo sapiens PAC cl	gb_pr:AC092782	+	8.00	93.65	7.5e+03	155355	AC092782 Homo sapiens chr
gb_pr:HS273H14	+	8.00	97.11	4.8e+03	95556	AL0808272	Human DNA sequence	gb_htg:AC015953	+	8.00	93.62	7.5e+03	155904	AC015953 Homo sapiens chr
gb_htg:AC016004	+	8.00	97.06	4.8e+03	96251	AC016004	Homo sapiens clone R	gb_in:AC023748	+	8.00	93.61	7.5e+03	156509	AC023748 Drosophila melan
gb_pr:HS633020	+	8.00	97.05	4.8e+03	96460	AL022804	Human DNA sequence	gb_pr:AC021914	+	8.00	93.58	7.5e+03	156507	AC021914 Homo sapiens chr
gb_htg:AC106033	-	8.00	97.03	4.8e+03	96732	AC106033	Homo sapiens Chromos	gb_pr:AC003396	+	8.00	93.57	7.5e+03	156829	AC003396 Homo sapiens chr
gb_htg:AC095539	+	8.00	96.92	4.9e+03	97307	AC095539	Rattus norvegicus cl	gb_pr:AL450333	+	8.00	93.57	7.5e+03	156964	AL450333 Human DNA sequen
gb_htg:AC096236	-	8.00	96.91	4.9e+03	98211	AC096236	Rattus norvegicus cl	gb_htg:AF1670729	+	8.00	93.57	7.5e+03	157033	AF1670729 Homo sapiens chr
gb_htg:AP000665	-	8.00	96.91	4.9e+03	98305	AP000665	Homo sapiens Chromos	gb_htg:AC087109	+	8.00	93.57	7.5e+03	157192	AC087109 Homo sapiens chr
gb_htg:AL592068	+	8.00	96.90	4.9e+03	98454	AL592068	Homo sapiens Chromos	gb_pr:AC092670	+	8.00	93.53	7.6e+03	157951	AC092670 Homo sapiens chr
gb_pr:AP000099	+	8.00	96.79	5.0e+03	100000	AP000099	Homo sapiens genom	gb_pr:AC011155	+	8.00	93.51	7.6e+03	158371	AC011155 Homo sapiens, cl
gb_pr:AP000175	+	8.00	96.79	5.0e+03	100000	AP000175	Homo sapiens genom	gb_htg:AC091404	+	8.00	93.49	7.6e+03	158755	AC091404 Sus scrofa clon
gb_pr:AP003087	+	8.00	96.76	5.0e+03	100403	AP003087	Homo sapiens genom	gb_pr:AC006013	+	8.00	93.48	7.6e+03	159002	AC006013 Homo sapiens PAC
gb_pr:AL390066	+	8.00	96.54	5.2e+03	103611	AL390066	Human DNA sequence	gb_htg:AC068293	+	8.00	93.47	7.6e+03	159339	AC068293 Homo sapiens chr
gb_htg:AC014357	+	8.00	96.53	5.2e+03	103746	AC014357	Drosophila melanogaster	gb_pr:AC074008	+	8.00	93.46	7.6e+03	159439	AC074008 Homo sapiens BAC
gb_htg:AC094679	+	8.00	96.27	5.3e+03	107637	AC094679	Rattus norvegicus cl	gb_pr:AC022919	+	8.00	93.45	7.7e+03	159749	AC022919 Homo sapiens chr
gb_pr:AC099783	+	8.00	96.21	5.4e+03	108400	L77570	Homo sapiens Digeorg	gb_htg:AC099783	+	8.00	93.44	7.7e+03	159870	AC099783 Homo sapiens chr
gb_pr:AC026736	+	8.00	96.20	5.4e+03	108661	AC026736	Homo sapiens Chromo	gb_pr:AC009082	+	8.00	93.42	7.7e+03	160277	AC009082 Homo sapiens chr
gb_pr:AC083949	-	8.00	96.16	5.4e+03	110096	AC083949	Homo sapiens Chromo	gb_htg:AC096203	+	8.00	93.42	7.7e+03	160323	AC096203 Rattus norvegicus
gb_htg:AC099666_1	-	8.00	96.11	5.4e+03	110000	Continuation (2 of 4) of AC0		gb_htg:AC098166	+	8.00	93.42	7.7e+03	160424	AC098166 Rattus norvegicus
gb_htg:AL627224_1	+	8.00	96.11	5.4e+03	110000	Continuation (2 of 5) of AL6		gb_htg:AC068372	+	8.00	93.42	7.7e+03	160426	AC068372 Homo sapiens chr

gb_hhg:AC021933	+	8.00	93.40	7.7e+03	160864	AC021933	Homo sapiens chromd	gb_pr:AC092724	+	8.00	92.73	8.4e+03	176521	AC092724	Homo sapiens chr
gb_pr:AC010138	-	8.00	93.39	7.7e+03	160970	AC010138	Homo sapiens chromd	gb_in:AC008029	+	8.00	92.72	8.4e+03	176991	AC008029	Drosophila melan
gb_hhg:AC087464	-	8.00	93.38	7.7e+03	161262	AC087464	Homo sapiens chromd	gb_pr:AF192304	+	8.00	92.71	8.4e+03	179226	AF192304	Homo sapiens chr
gb_pr:HS169M13	+	8.00	93.37	7.7e+03	161273	AL035427	Human DNA sequence	gb_hhg:AC096933	-	8.00	92.70	8.4e+03	177276	AC096933	Rattus norvegicus
gb_hhg:AC024229	-	8.00	93.37	7.7e+03	161540	AC024229	Homo sapiens chromd	gb_hhg:AL626782	-	8.00	92.70	8.4e+03	177475	AL626782	Mus musculus chr
gb_hhg:AC010514	-	8.00	93.37	7.7e+03	161542	AC0106148	Rattus norvegicus c	gb_hhg:AC098062	-	8.00	92.70	8.4e+03	177509	AC098062	Rattus norvegicus
gb_pr:AL1392045	+	8.00	93.35	7.8e+03	161933	AL1392045	Human DNA sequence	gb_hhg:AC026906	-	8.00	92.68	8.5e+03	177904	AC026906	Homo sapiens chr
gb_hhg:AC002986	+	8.00	93.34	7.8e+03	162066	AP002986	Homo sapiens chromd	gb_hhg:AC096936	-	8.00	92.67	8.5e+03	178073	AC096936	Rattus norvegicus
gb_hhg:AC019271	+	8.00	93.34	7.8e+03	162167	AC019271	Homo sapiens clone	gb_pr:AL590723	+	8.00	92.67	8.5e+03	178223	AL590723	Human DNA sequen
gb_hhg:AC009367	+	8.00	93.31	7.8e+03	162851	AC009637	Homo sapiens chromd	gb_hhg:AC009420	-	8.00	92.67	8.5e+03	178229	AC009420	Homo sapiens clo
gb_hhg:AC091925	-	8.00	93.31	7.8e+03	162914	AC091925	Homo sapiens chromd	gb_hhg:AC0011968	-	8.00	92.66	8.5e+03	178239	AC0011968	Homo sapiens chr
gb_hhg:AC078804	+	8.00	93.29	7.8e+03	163243	AC078804	Homo sapiens chromd	gb_hhg:AL645822	-	8.00	92.66	8.5e+03	178313	AL645822	Danio rerio clon
gb_pr:AC093868	+	8.00	93.26	7.8e+03	163334	AC093868	Homo sapiens chromd	gb_hhg:AC016383	+	8.00	92.65	8.5e+03	178680	AC016383	AC016383
gb_hhg:AC012132	+	8.00	93.26	7.8e+03	163990	AC012132	Homo sapiens clone	gb_pr:AC079316	+	8.00	92.64	8.5e+03	178905	AC079316	Homo sapiens chr
gb_pr:AC068614	+	8.00	93.26	7.9e+03	164059	AC068614	Rattus norvegicus c	gb_hhg:AP001934	+	8.00	92.62	8.5e+03	179281	AP001934	AP001934
gb_hhg:AC026657	+	8.00	93.23	7.9e+03	164706	AC026657	Homo sapiens BAC cl	gb_hhg:AC0090107	+	8.00	92.62	8.5e+03	179470	AC0090107	Homo sapiens chr
gb_pr:CNS010UN	-	8.00	93.23	7.9e+03	164766	AC026657	Homo sapiens chromd	gb_hhg:AC0022177	+	8.00	92.59	8.6e+03	180076	AC022177	Homo sapiens clo
gb_hhg:AC092729	-	8.00	93.22	7.9e+03	164935	AL133329	Human chromosome 14	gb_hhg:AC104820	+	8.00	92.58	8.6e+03	180395	AC104820	Homo sapiens chr
gb_pr:CNS01DRF	-	8.00	93.22	7.9e+03	164958	AC092729	Human chromosome 14	gb_pr:CNS01DRF	-	8.00	92.57	8.6e+03	180523	AC092729	Human chromosome
gb_hhg:AC092729	-	8.00	93.21	7.9e+03	165049	AL117190	Human chromosome 14	gb_hhg:AC022080	-	8.00	92.57	8.6e+03	180676	AC022080	Homo sapiens chr
gb_pr:CNS01DRF	-	8.00	93.21	7.9e+03	165159	AL133454	Human chromosome 14	gb_hhg:AC104299	+	8.00	92.57	8.6e+03	180725	AC104299	Homo sapiens chr
gb_hhg:AC026180	+	8.00	93.16	8.0e+03	166352	AC026180	Homo sapiens chromd	gb_hhg:AC092250	+	8.00	92.55	8.6e+03	181119	AC092250	Homo sapiens chr
gb_pr:AP002783	+	8.00	93.14	8.0e+03	166772	AC098691	Homo sapiens chromd	gb_hhg:AC097538	-	8.00	92.54	8.6e+03	181525	AC097538	Rattus norvegicus
gb_hhg:AC000750	+	8.00	93.13	8.0e+03	167026	AP002783	Homo sapiens genom	gb_pr:AF411058	+	8.00	92.53	8.6e+03	181654	AF411058	Homo sapiens clo
gb_hhg:AC103484	+	8.00	93.10	8.0e+03	167498	AC095510	Homo sapiens chromd	gb_pr:AC097714	+	8.00	92.53	8.6e+03	181756	AC097714	Homo sapiens chr
gb_hhg:AC025094	+	8.00	93.07	8.0e+03	167600	AC095510	Rattus norvegicus c	gb_hhg:AC103954	+	8.00	92.52	8.6e+03	181904	AC103954	Homo sapiens chr
gb_hhg:AC058792	-	8.00	93.06	8.0e+03	168481	AC025094	Homo sapiens chromd	gb_hhg:AL535872	+	8.00	92.52	8.6e+03	182042	AL535872	Homo sapiens chr
gb_hhg:AC022954	-	8.00	93.06	8.0e+03	168540	AC058792	Homo sapiens chromd	gb_hhg:AC021188	+	8.00	92.51	8.6e+03	182106	AC021188	Homo sapiens chr
gb_pr:AC0210854	+	8.00	93.02	8.1e+03	169370	AC022954	Homo sapiens clone	gb_hhg:AC021128	+	8.00	92.51	8.6e+03	182267	AC021128	Homo sapiens chr
gb_hhg:AC091147	+	8.00	93.01	8.1e+03	169836	AC010854	Homo sapiens chromd	gb_hhg:AC023088	+	8.00	92.48	8.7e+03	182600	AC023088	Homo sapiens chr
gb_hhg:AC0902435	+	8.00	93.00	8.1e+03	170028	AC091147	Homo sapiens clone	gb_pr:AL162578	+	8.00	92.48	8.7e+03	182823	AL162578	Human DNA sequen
gb_hhg:AC018371	+	8.00	93.00	8.1e+03	170119	AP002435	Homo sapiens chromd	gb_hhg:AC106518	+	8.00	92.46	8.7e+03	183553	AC106518	Rattus norvegicus
gb_hhg:AC000773	+	8.00	92.99	8.1e+03	170133	AC091870	Homo sapiens chromd	gb_hhg:AC106545	+	8.00	92.45	8.7e+03	183646	AC106545	Rattus norvegicus
gb_pr:AC000773	+	8.00	92.99	8.1e+03	170204	AC018371	Homo sapiens chromd	gb_pr:AL161646	+	8.00	92.45	8.7e+03	183833	AL161646	Human DNA sequen
gb_hhg:AC0094970	+	8.00	92.98	8.1e+03	170634	AC018539	Homo sapiens chromd	gb_hhg:AC018708	+	8.00	92.42	8.7e+03	184516	AC018708	Homo sapiens chr
gb_hhg:AC023429	+	8.00	92.97	8.1e+03	170682	AP000773	Homo sapiens genom	gb_pr:AC096242	+	8.00	92.41	8.7e+03	184664	AC096242	Homo sapiens chr
gb_pr:AC087456	+	8.00	92.96	8.2e+03	170945	AC023429	Homo sapiens chromd	gb_hhg:AP000849	+	8.00	92.41	8.7e+03	184794	AP000849	Homo sapiens chr
gb_hhg:AC004382	+	8.00	92.95	8.2e+03	171321	AC087456	Homo sapiens chromd	gb_pr:AC099314	+	8.00	92.40	8.8e+03	184945	AC099314	Homo sapiens chr
gb_hhg:AC0091541	+	8.00	92.93	8.2e+03	171581	AC043832	Homo sapiens chromd	gb_hhg:AC096641	+	8.00	92.39	8.8e+03	185148	AC096641	Homo sapiens chr
gb_hhg:AC009278	+	8.00	92.93	8.2e+03	171851	AC023287	Homo sapiens chromd	gb_hhg:AC006508	+	8.00	92.39	8.8e+03	185286	AC006508	Mus musculus Yp
gb_hhg:AC0090768	-	8.00	92.93	8.2e+03	171871	AC091541	Canis familiaris cl	gb_hhg:AL591723	+	8.00	92.37	8.8e+03	185644	AL591723	Homo sapiens chr
gb_hhg:AC0091419	-	8.00	92.93	8.2e+03	171890	AC093328	Homo sapiens chromd	gb_hhg:AC087163	+	8.00	92.37	8.8e+03	185686	AC087163	Homo sapiens chr
gb_hhg:AC001176	-	8.00	92.92	8.2e+03	171930	AC090768	Homo sapiens chromd	gb_hhg:AL645647	+	8.00	92.35	8.8e+03	186278	AL645647	Mus musculus chr
gb_hhg:AC0010575	+	8.00	92.91	8.2e+03	172182	AC097419	Rattus norvegicus c	gb_hhg:AL670169	+	8.00	92.34	8.8e+03	186497	AL670169	Homo sapiens chr
gb_hhg:AC0092331	-	8.00	92.90	8.2e+03	172401	AC025130	Homo sapiens chromd	gb_hhg:AL671978	+	8.00	92.31	8.9e+03	187274	AL671978	Homo sapiens chr
gb_hhg:AC0098117	+	8.00	92.89	8.2e+03	172592	AC079816	Mus musculus clone	gb_hhg:AL590410	+	8.00	92.26	8.9e+03	188538	AL590410	Drosophila melan
gb_hhg:AC0025130	+	8.00	92.88	8.2e+03	172837	AC010575	Drosophila melanog	gb_hhg:AL671978	+	8.00	92.26	8.9e+03	188615	AL671978	Mus musculus chr
gb_hhg:AC004359	-	8.00	92.88	8.2e+03	173052	AC099231	Rattus norvegicus c	gb_hhg:AL671978	+	8.00	92.26	8.9e+03	188615	AL671978	Mus musculus chr
gb_hhg:AC0034272	+	8.00	92.86	8.3e+03	173422	AC098117	Rattus norvegicus c	gb_hhg:AL671978	+	8.00	92.26	8.9e+03	188615	AL671978	Mus musculus chr
gb_hhg:AC007467	+	8.00	92.86	8.3e+03	173479	AC025130	Homo sapiens chromd	gb_hhg:AL671978	+	8.00	92.26	8.9e+03	188615	AL671978	Mus musculus chr
gb_hhg:AC009520	+	8.00	92.86	8.3e+03	173479	AC093568	Oryza sativa chromd	gb_hhg:AL671978	+	8.00	92.26	8.9e+03	188615	AL671978	Mus musculus chr
gb_hhg:AC0021145	+	8.00	92.86	8.3e+03	173484	AC093568	Oryza sativa chromd	gb_hhg:AL671978	+	8.00	92.26	8.9e+03	188615	AL671978	Mus musculus chr
gb_pr:AC093568	+	8.00	92.85	8.3e+03	173644	AC093568	Oryza sativa chromd	gb_hhg:AL671978	+	8.00	92.26	8.9e+03	188615	AL671978	Mus musculus chr
gb_pr:AL626785	-	8.00	92.82	8.3e+03	174098	AC005737	Homo sapiens chromd	gb_hhg:AC073389	+	8.00	92.01	9.2e+03	194466	AC073389	Homo sapiens chr
gb_hhg:AC102971	+	8.00	92.82	8.3e+03	174365	AL626785	Mus musculus chromd	gb_hhg:AC032024	+	8.00	92.01	9.2e+03	194466	AC032024	Homo sapiens chr
gb_pr:AP002342	-	8.00	92.82	8.3e+03	174548	AC102971	Rattus norvegicus c	gb_hhg:AC087732	+	8.00	92.00	9.2e+03	195599	AC087732	Homo sapiens chr
gb_hhg:AC0023889	+	8.00	92.81	8.3e+03	174612	AP002342	Homo sapiens genom	gb_hhg:AL535651	+	8.00	91.97	9.3e+03	196414	AL535651	Homo sapiens chr
gb_hhg:AC0021490	-	8.00	92.81	8.3e+03	174637	AC023889	Homo sapiens chromd	gb_pr:AL535825	+	8.00	91.89	9.3e+03	196794	AL535825	Human DNA sequen
gb_hhg:AC0016896	-	8.00	92.81	8.3e+03	174742	AC027490	Homo sapiens chromd	gb_hhg:AL645668	+	8.00	91.88	9.4e+03	199008	AL645668	Mus musculus chr
gb_hhg:AC0010274	-	8.00	92.79	8.3e+03	175071	AC016896	Homo sapiens chromd	gb_hhg:AL645668	+	8.00	91.88	9.4e+03	199008	AL645668	Mus musculus chr
gb_hhg:AC0098121	+	8.00	92.79	8.3e+03	175218	AC010274	Homo sapiens chromd	gb_hhg:AL645668	+	8.00	91.87	9.4e+03	199261	AL645668	Mus musculus chr
gb_hhg:AC0105657	+	8.00	92.77	8.4e+03	175660	AC097821	Rattus norvegicus c	gb_pr:ATGCRIV45	+	8.00	91.85	9.4e+03	199548	ATGCRIV45	Aradidopsis thal
gb_hhg:AC0095485	+	8.00	92.76	8.4e+03	175944	AC105657	Rattus norvegicus c	gb_hhg:AC095457	+	8.00	91.85	9.4e+03	199708	AC095457	Rattus norvegicus
gb_hhg:AC0091100	-	8.00	92.75	8.4e+03	176233	AC095485	Rattus norvegicus c	gb_pr:ATGCRIV45	+	8.00	91.77	9.5e+03	202050	ATGCRIV45	Aradidopsis thal
gb_hhg:AC0024418	+	8.00	92.74	8.4e+03	176425	AC091100	Homo sapiens chromd	gb_hhg:AL607108	-	8.00	91.74	9.5e+03	202961	AL607108	Rattus norvegicus
						AC024449	Homo sapiens chromd	gb_hhg:AL607108	-	8.00	91.74	9.5e+03	202961	AL607108	Rattus norvegicus
						AC090418	Homo sapiens chromd	gb_hhg:AC006342	+	8.00	91.74	9.5e+03	203046	AC006342	Homo sapiens chr

gb.com:AF162514	+	7.00	115.65	444.57	462	AF162514 Bos grunniens satellite	gb.pr:HSN25321	-	7.00	111.91	718.00	780	AJ25321 Homo sapiens genomic
gb.pl:TL047411	-	7.00	115.61	446.33	464	U47411 Rattus norvegicus histon	em.hito.inh:AC056122	-	7.00	111.91	718.00	780	AC056122 Giardia intestinalis
gb.pr:HSR9GVN	-	7.00	115.48	454.25	473	X05504 Human rearranged gene TR	em.hito.inh:AC074653	-	7.00	111.88	719.69	782	AC074653 Giardia intestinalis
gb.pr:HSR9GVN	-	7.00	115.46	455.13	474	X08086 Human rearranged gene TR	gb.pr:HSN335561	-	7.00	111.88	720.53	783	AJ33561 Homo sapiens genomic
gb.sts:HS006611	+	7.00	115.40	458.64	478	M57470 Murine beta-galactosidase	em.hito.inh:AC030174	-	7.00	111.73	734.83	800	AC030174 Giardia intestinalis
gb.pl:IDV47383	-	7.00	115.24	468.29	489	U47383 Drosophila histone	em.hito.inh:AC081799	-	7.00	111.63	744.08	811	AC081799 Giardia intestinalis
gb.ro:MM44KDE1	-	7.00	115.15	472.68	494	U47383 Drosophila histone	gb.pl:YSPD2305	+	7.00	111.59	748.27	816	AJ22305 Sequence 2221 from
gb.ro:MM44KDE1	-	7.00	115.15	473.55	495	X53067 Mouse mRNA for 14kDa lec	em.hito.inh:AC071100	+	7.00	111.53	753.31	822	L07304 Schistosoma mansoni p
gb.pac:A27894	-	7.00	115.12	473.55	495	A27894 Coding sequence for GGP	em.hito.inh:AC071072	+	7.00	111.51	755.82	825	AC071072 Giardia intestinalis
gb.pac:A27894	-	7.00	115.12	473.55	497	A27894 Coding sequence for GGP	em.hito.inh:AC050107	+	7.00	111.45	761.69	832	AC050107 Giardia intestinalis
gb.pac:A27894	-	7.00	115.12	473.55	497	A27894 Coding sequence for GGP	gb.pac:A68691	+	7.00	111.41	765.04	836	A68691 Sequence 5 from Paten
gb.sts:AF046519	-	7.00	115.11	476.18	498	AF046519 Rattus norvegicus, ORF	em.hito.inh:AC089248	+	7.00	111.35	771.74	844	AC089248 Giardia intestinalis
gb.com:AF045515	+	7.00	115.10	477.05	499	AF045515 Sus scrofa breed large	em.hito.inh:AC034448	+	7.00	111.33	773.41	846	AC034448 Giardia intestinalis
gb.ba:AB039813	-	7.00	115.02	481.43	504	AB039817 Chlorella sp. HS-7	em.hito.inh:AC071100	-	7.00	111.29	776.76	850	AC071100 Giardia intestinalis
gb.pl:AF114969	-	7.00	114.86	491.91	516	AF114969 Schistosoma castellan	gb.pr:HSR9GVN	-	7.00	111.28	778.43	852	X71288 Rattus norvegicus p
gb.ro:AF095934	-	7.00	114.67	504.11	530	AF095934 Dictyostellum discoide	gb.pl:JUN278454	-	7.00	111.24	781.78	856	JUN278454 Juncus nigra mRNA
gb.pr:HSR9GVN	+	7.00	114.67	504.11	530	AF073345 Erythrina chloroacetam	gb.pac:A85851	+	7.00	111.12	794.31	871	A85851 Sequence 510 from Pat
gb.pac:A8070588	-	7.00	114.65	504.98	531	AR070588 Sequence 19 from Paten	gb.pac:A85851	+	7.00	111.12	794.31	871	A85851 Sequence 510 from Pat
gb.pac:A8070588	-	7.00	114.59	509.33	536	AR070588 Sequence 19 from Paten	em.hito.inh:AC031263	+	7.00	111.12	794.31	871	AC031263 Giardia intestinalis
gb.ro:BC002063	-	7.00	114.47	511.94	539	BC002063 Mus musculus, lectin	em.hito.inh:AC044285	+	7.00	111.09	797.65	875	AC044285 Giardia intestinalis
gb.ro:MM141L	-	7.00	114.45	517.15	545	G65523 Mus musculus mRNA for IL4	gb.pac:A805502	+	7.00	110.98	808.45	888	AC05502 Sequence 1 from Pat
gb.pl:SOBETUA	-	7.00	114.43	519.76	548	L47262 Sportidololus paratensis	gb.pr:BC007282	+	7.00	110.97	809.32	889	BC007282 Homo sapiens, clone
gb.ov:AF297036	+	7.00	114.36	524.10	553	AF297036 Ophiophagus hannah ubi	em.hito.inh:AC028084	+	7.00	110.87	820.14	902	AF228242 Plagioplecton suaveo
gb.pl:KB069580	-	7.00	114.20	535.36	566	U69580 Kluyveromyces fragilis 26	gb.pl:AF222424	-	7.00	110.85	822.64	905	AF222424 Plagioplecton suaveo
gb.pr:HSN33561	-	7.00	114.18	536.23	567	M63999 Trilicium aestivum DNA b1	em.hito.inh:AC058409	-	7.00	110.80	827.63	911	AC058409 Giardia intestinalis
gb.pr:HSN33561	-	7.00	114.16	537.96	569	AJ331882 Homo sapiens genomic s	gb.sts:AB030249	-	7.00	110.78	828.46	912	AB030249 Synthetic Mus muscu
gb.sts:G56256	+	7.00	114.15	538.82	570	G56256 SHGC-101770 Human Homo s	em.hito.inh:AC053782	+	7.00	110.78	830.12	914	AC053782 Giardia intestinalis
gb.in:AF165428	-	7.00	114.03	546.60	579	AF165428 Aphus nerii esterase p	em.hito.inh:AC084990	+	7.00	110.75	833.43	918	AC084990 Giardia intestinalis
gb.pac:AJ330046	-	7.00	114.03	546.60	579	AF165428 Aphus nerii esterase p	em.hito.inh:AC084990	+	7.00	110.75	833.43	918	AC084990 Giardia intestinalis
gb.pac:AJ330046	-	7.00	114.01	548.33	581	AF165428 Aphus nerii esterase p	em.hito.inh:AC084990	+	7.00	110.75	833.43	918	AC084990 Giardia intestinalis
gb.sts:G58481	-	7.00	114.01	548.33	581	G58481 SHGC-10443 Human Homo s	gb.ov:CNM11396	-	7.00	110.68	840.93	927	AC058409 Giardia intestinalis
gb.sts:G58481	-	7.00	113.96	551.79	585	G48759 SHGC-79126 Human Homo s	em.hito.inh:AC053110	-	7.00	110.54	855.86	945	AC053110 Giardia intestinalis
gb.pr:HMWCGXK	-	7.00	113.77	565.58	601	AL158732 Homo sapiens T cell rece	gb.pr:HSN330778	+	7.00	110.51	859.17	949	AJ330778 Homo sapiens genomic
gb.sts:HS07661	+	7.00	113.76	566.45	602	AC039463 Giardia intestinalis c	em.hito.inh:AC057823	+	7.00	110.50	860.00	950	AC057823 Giardia intestinalis
em.hito.inh:AC039463	+	7.00	113.76	566.45	602	AC039463 Giardia intestinalis c	gb.com:AF080194	-	7.00	110.46	864.97	956	AF080194 Smuttophys murina
em.hito.inh:AC039463	+	7.00	113.60	572.47	609	AC039463 Giardia intestinalis c	em.hito.inh:AC089665	-	7.00	110.45	865.80	957	AF089665 Smuttophys murina
em.hito.inh:AC039463	+	7.00	113.60	572.47	609	AC039463 Giardia intestinalis c	em.hito.inh:AC089665	-	7.00	110.45	865.80	957	AF089665 Smuttophys murina
gb.pr:HSN330402	+	7.00	113.52	583.65	615	AC039463 Giardia intestinalis c	em.hito.inh:AC076113	-	7.00	110.44	866.63	958	AC076113 Giardia intestinalis
gb.ro:MM68B93	-	7.00	113.32	589.09	640	X51578 M.musculus bdbp gene, ex	em.hito.inh:AC080849	+	7.00	110.44	866.63	958	AC080849 Giardia intestinalis
gb.ov:AY009126	+	7.00	113.08	617.91	662	AY009126 Oncorhynchus mykiss ve	gb.pl:AF21156	-	7.00	110.38	872.25	966	AF21156 Arabidopsis thalian
gb.ro:MUSCICAM04	-	7.00	113.05	620.47	665	M90545 Mouse intercellular adhe	gb.pr:AB060229	-	7.00	110.34	878.21	972	AB060229 Macaca fasciculari
gb.ov:AF273369	-	7.00	113.04	621.32	666	AF273369 Xenopus sp. clone XE2	em.hito.inh:AC081621	-	7.00	110.29	883.17	978	AC081621 Giardia intestinalis
gb.ov:AF273369	-	7.00	112.98	629.55	671	AF273369 Xenopus sp. clone XE2	em.hito.inh:AC031105	-	7.00	110.29	884.00	979	AC031105 Giardia intestinalis
gb.pr:HSN336244	-	7.00	112.93	629.55	676	AJ262444 Homo sapiens genomic s	gb.ba:AF142842	-	7.00	110.24	888.96	985	AF142842 Uncultured bacteriu
gb.ba:U8A309618	+	7.00	112.91	631.56	678	AJ262444 Homo sapiens genomic s	em.hito.inh:AC076538	-	7.00	110.20	893.91	991	AC076538 Giardia intestinalis
gb.pr:HSN3329640	+	7.00	112.88	634.12	681	AJ262444 Homo sapiens genomic s	em.hito.inh:AC053534	-	7.00	110.16	898.04	996	AC053534 Giardia intestinalis
gb.ba:SHRTRP	-	7.00	112.86	635.82	683	J01787 Shigella dysenteriae try	gb.pl:WHHBPAL	-	7.00	110.16	898.04	997	DI2919 Trilicium aestivum mRN
gb.pr:HSN3329640	-	7.00	112.86	635.82	683	J01787 Shigella dysenteriae try	gb.pr:HSN3329640	-	7.00	110.14	900.51	999	AJ332964 Homo sapiens genomic
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29	700	AX182416 Sequence 169 from Pat	em.hito.inh:AC054312	+	7.00	110.13	902.16	1001	AC054312 Giardia intestinalis
gb.pac:AJ182416	-	7.00	112.68	650.29									

gb_vt:EHU43560	-	7.00	109.08	1.0e+03	1159	U43560 Epizootic hemorrhagic d	gb_pat:AX061643	-	7.00	106.68	1.4e+03	1622	AX061643 Sequence 48 from P
gb_vt:AB002869	+	7.00	109.06	1.0e+03	1162	AB002869 Subacute sclerosing m	gb_vt:OYE22223	-	7.00	106.67	1.4e+03	1625	AJ292223 Onion yellow dwarf
gb_vt:AB016239	-	7.00	109.06	1.0e+03	1162	AB016239 Measles virus gene f	gb_ov:AF124511	-	7.00	106.65	1.4e+03	1629	AF124511 Gallus gallus BVES
gb_vt:HSR8399F	-	7.00	108.95	1.0e+03	1180	X15274 H.sapiens TRG9 gene, a	gb_vt:G73185	-	7.00	106.65	1.4e+03	1629	G73185 csprad1-pcr2-4 Huma
gb_vt:HSR279141	-	7.00	108.94	1.1e+03	1183	AJ279141 Homo sapiens partial	gb_vt:AB070034	-	7.00	106.60	1.4e+03	1640	AB070034 Macaca fasciculari
gb_vt:BC008032	-	7.00	108.92	1.1e+03	1185	BC008032 Homo sapiens, clone 1	gb_vt:BC011575	-	7.00	106.60	1.4e+03	1641	BC011575 Homo sapiens, Simi
gb_vt:HSU29441	-	7.00	108.90	1.1e+03	1189	U29441 Diplomonad ATCC50330 be	gb_vt:PSA10CEB1	-	7.00	106.60	1.4e+03	1642	Y13328 Pichia stipitis ADH1
gb_vt:AY051038	-	7.00	108.85	1.1e+03	1198	AY051038 Arabidopsis thaliana	gb_vt:BC028284	-	7.00	106.55	1.4e+03	1653	BC028284 Homo sapiens, seri
gb_vt:PAU27709	-	7.00	108.82	1.1e+03	1203	U27709 Pseudomonas aeruginosa	gb_vt:AF370135	-	7.00	106.54	1.4e+03	1655	AF370135 Arabidopsis thalia
gb_vt:EPHEP1AD	-	7.00	108.82	1.1e+03	1203	D49925 Ephydratia fluviatilis m	gb_vt:AY042904	-	7.00	106.52	1.4e+03	1659	AY042904 Arabidopsis thalia
gb_vt:AX083327	-	7.00	108.73	1.1e+03	1218	AX083327 Sequence 19 from Pat	gb_vt:AX06562	-	7.00	106.51	1.4e+03	1662	AX06562 Sequence 595 from Pa
gb_vt:AX002986	-	7.00	108.57	1.1e+03	1245	AX002986 Sequence 5 from Paten	gb_vt:AX014962	-	7.00	106.48	1.4e+03	1669	AX014962 Sequence 8 from Pa
gb_vt:AF132515	-	7.00	108.55	1.1e+03	1249	AF132515 Lutzmomyia longipalpis	gb_vt:AX050448	-	7.00	106.48	1.4e+03	1669	AX050448 Sequence 8 from Pa
gb_vt:AY069764	-	7.00	108.51	1.1e+03	1256	AY069764 Drosophila melanogast	gb_vt:AX336106	-	7.00	106.48	1.4e+03	1670	AX336106 Sequence 6615 from
gb_vt:LE5297916	-	7.00	108.48	1.1e+03	1261	AJ297916 Lycopersicon esculent	gb_vt:AX336460	-	7.00	106.48	1.4e+03	1669	AX336460 Sequence 6969 from
gb_vt:1HX18854	-	7.00	108.43	1.1e+03	1270	Y18854 Infectious hematopoiet	gb_vt:HMUSHMTA	-	7.00	106.46	1.4e+03	1674	HMUSHMTA
gb_vt:AF139500	-	7.00	108.37	1.1e+03	1280	AF139500 Prunus americana p9f	gb_vt:AX0300175	-	7.00	106.46	1.4e+03	1674	AX0300175 Sequence 1 from Pa
gb_vt:AK055285	-	7.00	108.37	1.1e+03	1281	AK055285 Homo sapiens, cDNA FL	gb_vt:AS6568	-	7.00	106.45	1.4e+03	1677	AS6568 Sequence 597 from Pa
gb_vt:BC013062	-	7.00	108.36	1.1e+03	1282	BC013062 Homo sapiens, Simila	gb_vt:AS6568	-	7.00	106.45	1.4e+03	1677	AS6568 Sequence 601 from Pa
gb_vt:MUSELPL2	-	7.00	108.33	1.1e+03	1288	D49682 Mouse FT2-F1 gene, 2/19	gb_vt:MUSADPP	-	7.00	106.43	1.4e+03	1680	M93375 Mouse adipose differ
gb_vt:AF000204	-	7.00	108.29	1.1e+03	1295	AF000204 Rattus musculus ADP-rib	gb_vt:AX149469	-	7.00	106.42	1.5e+03	1683	AX149469 Sequence 25 from P
gb_vt:AF000204	-	7.00	108.28	1.1e+03	1297	U87600 Rattus norvegicus L1 re	gb_vt:123946	-	7.00	106.41	1.5e+03	1685	123946 Sequence 1 from pate
gb_vt:RN1LMY11	-	7.00	108.28	1.1e+03	1297	U87600 Rattus norvegicus L1 re	gb_vt:AX051970	-	7.00	106.39	1.5e+03	1690	AX051970 Sequence 24 from p
gb_vt:153503	-	7.00	108.27	1.1e+03	1299	153503 Sequence 3 from patent	gb_vt:152144	-	7.00	106.39	1.5e+03	1690	152144 Sequence 24 from pat
gb_vt:AF000199	-	7.00	108.23	1.2e+03	1306	AF000199 Rattus sp. T-612 retr	gb_vt:AX000050	-	7.00	106.39	1.5e+03	1690	AX000050 Homo sapiens cDNA
gb_vt:HS11POBPL	-	7.00	108.19	1.2e+03	1314	X97324 H.sapiens mRNA for adif	gb_vt:AF242529	-	7.00	106.37	1.5e+03	1695	AF242529 Homo sapiens prote
gb_vt:AF443203	-	7.00	108.19	1.2e+03	1314	AF443203 Homo sapiens adipose	gb_vt:1579359	-	7.00	106.35	1.5e+03	1700	1579359 arabidogalactan prot
gb_vt:AY038335	-	7.00	108.13	1.2e+03	1325	AY038335 Inocbe sp. Triappe 25	gb_vt:AF226646	-	7.00	106.29	1.5e+03	1714	AF226646 Turnip mosaic viru
gb_vt:AT271731	-	7.00	108.09	1.2e+03	1332	AJ271731 Arabidopsis thaliana	gb_vt:1DC0470748	-	7.00	106.27	1.5e+03	1719	1DC0470748
gb_vt:AX149467	-	7.00	108.01	1.2e+03	1347	AX149467 Sequence 23 from Pat	gb_vt:AC106945	-	7.00	106.25	1.5e+03	1723	AC106945 Rattus norvegicus
gb_vt:184655	-	7.00	107.98	1.2e+03	1352	184655 Sequence 10 from patent	gb_vt:AB049055	-	7.00	106.23	1.5e+03	1736	AB049055 Mus musculus Brail
gb_vt:AY061083	-	7.00	107.87	1.2e+03	1373	AY061083 Drosophila melanogast	gb_vt:AB045893	-	7.00	106.13	1.5e+03	1753	AB045893 Nephenthes alata Na
gb_vt:AY069455	-	7.00	107.85	1.2e+03	1376	AY069455 Drosophila melanogast	gb_vt:HV0310848	-	7.00	106.01	1.5e+03	1782	AF130848 Hordeum vulgare pa
gb_vt:HSN32592	-	7.00	107.84	1.2e+03	1377	AJ325792 Homo sapiens genomic	gb_vt:AF197904	-	7.00	105.99	1.5e+03	1787	AF197904 Saurus aureata para
gb_vt:SM80056	-	7.00	107.84	1.2e+03	1380	U80056 Amsacta moorei entomop	gb_vt:AF197904	-	7.00	105.99	1.5e+03	1787	AF197904 Saurus aureata para
gb_vt:WHREMGFE13	-	7.00	107.73	1.2e+03	1388	S69363 Protein C inhibitor (hu	gb_vt:AF130441	-	7.00	105.87	1.6e+03	1818	AF130441 Arabidopsis thalia
gb_vt:AB049056	-	7.00	107.70	1.2e+03	1406	AB049056 Rattus norvegicus bra	gb_vt:OC038195	-	7.00	105.86	1.6e+03	1821	OC038195 Oryzotolagus cunicul
gb_vt:AB050224	-	7.00	107.70	1.2e+03	1407	AB050224 Uncultured archaeon S	gb_vt:AX086523	-	7.00	105.75	1.6e+03	1841	AX086523 Sequence 475 from
gb_vt:AB050225	-	7.00	107.67	1.2e+03	1407	AB050225 Uncultured archaeon S	gb_vt:HSN801723	-	7.00	105.72	1.6e+03	1857	HSN801723
gb_vt:119715	-	7.00	107.67	1.2e+03	1412	119715 Sequence 1 from patent	gb_vt:RNB16	-	7.00	105.72	1.6e+03	1857	RNB16
gb_vt:MMU65021	-	7.00	107.67	1.2e+03	1413	AF039461 Mus musculus Leptin r	gb_vt:MMU65021	-	7.00	105.71	1.6e+03	1860	MMU65021 Mus musculus amelobl
gb_vt:TRNEHLEA	-	7.00	107.65	1.2e+03	1416	L20680 Thermus aquaticus DNA r	gb_vt:HS14485	-	7.00	105.68	1.6e+03	1867	HS14485 Homo sapiens mRNA fo
gb_vt:RNM1LTV11	-	7.00	107.55	1.2e+03	1429	U87593 Rattus norvegicus L1 re	gb_vt:BC006906	-	7.00	105.64	1.6e+03	1877	BC006906 Mus musculus, Simi
gb_vt:AF047384	-	7.00	107.56	1.3e+03	1435	AF047384 Rattus norvegicus pos	gb_vt:AX068976	-	7.00	105.55	1.6e+03	1890	AX068976 Sequence 1 from Pa
gb_vt:U38980	-	7.00	107.51	1.3e+03	1445	U38980 Human PMS2 related (PM	gb_vt:AY035782	-	7.00	105.55	1.6e+03	1900	AY035782 Rhizopcephalus appe
gb_vt:TA007933	-	7.00	107.25	1.3e+03	1499	U07933 Trifolium aestivum Chind	gb_vt:AB001278	-	7.00	105.55	1.6e+03	1900	AB001278 Sequence 22 from p
gb_vt:MM049107	-	7.00	107.23	1.3e+03	1502	U49107 Mus musculus Leptin re	gb_vt:AX025098	-	7.00	105.55	1.6e+03	1900	AX025098 Sequence 22 from p
gb_vt:AY061282	-	7.00	107.17	1.3e+03	1515	AY061282 Drosophila melanogast	gb_vt:AB001901	-	7.00	105.55	1.6e+03	1900	AB001901 Remedy of virus. 1
gb_vt:AX353444	-	7.00	107.16	1.3e+03	1518	AX353444 Sequence 10 from Pat	gb_vt:RNM21070	-	7.00	105.54	1.6e+03	1900	U21070 Rattus norvegicus GA
gb_vt:144674	-	7.00	107.15	1.3e+03	1520	144674 Sequence 1 from patent	gb_vt:AF056335	-	7.00	105.54	1.6e+03	1905	AF056335 Bacillus lichenifo
gb_vt:RNEHAB	-	7.00	107.15	1.3e+03	1520	Z28365 R.norvegicus BEHA mRNA	gb_vt:AK026783	-	7.00	105.52	1.6e+03	1914	AK026783 Homo sapiens cDNA:
gb_vt:AY054265	-	7.00	107.13	1.3e+03	1524	AY054265 Arabidopsis thaliana	gb_vt:AX038081	-	7.00	105.50	1.6e+03	1914	AX038081 Bombyx mori putati
gb_vt:AB001616	-	7.00	107.13	1.3e+03	1524	AB001616 Homo sapiens DNA for	gb_vt:BC005127	-	7.00	105.49	1.6e+03	1916	BC005127 Homo sapiens, adip
gb_vt:AF448803	-	7.00	107.09	1.3e+03	1532	AF448803 Human herpesvirus 4	gb_vt:TVCPPT	-	7.00	105.47	1.6e+03	1923	X56501 T.vulgaris cpi gene
gb_vt:AR023708	-	7.00	107.08	1.3e+03	1535	AR023708 Sequence 10 from pate	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40871 Infectious hematopoi
gb_vt:AR103242	-	7.00	107.01	1.3e+03	1550	AR103242 Sequence 3 from paten	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40872 Infectious hematopoi
gb_vt:BD000105	-	7.00	107.01	1.3e+03	1550	BD000105 Differentiation and b	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40873 Infectious hematopoi
gb_vt:HS058681	-	7.00	106.99	1.3e+03	1550	U58681 Homo sapiens neurogenic	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40874 Infectious hematopoi
gb_vt:AY065116	-	7.00	106.97	1.4e+03	1553	AY065116 Arabidopsis thaliana	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40875 Infectious hematopoi
gb_vt:AY060613	-	7.00	106.97	1.4e+03	1559	AY060613 Drosophila melanogast	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40876 Infectious hematopoi
gb_vt:AX136103	-	7.00	106.85	1.4e+03	1584	AX136103 Sequence 25 from Pat	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40877 Infectious hematopoi
gb_vt:AX349757	-	7.00	106.85	1.4e+03	1584	AX349757 Sequence 280 from Pat	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40878 Infectious hematopoi
gb_vt:AF038193	-	7.00	106.85	1.4e+03	1585	AF038193 Homo sapiens clone 23	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40879 Infectious hematopoi
gb_vt:HS14487	-	7.00	106.84	1.4e+03	1586	Y14487 Homo sapiens mRNA for	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40880 Infectious hematopoi
gb_vt:HMUSH	-	7.00	106.79	1.4e+03	1599	L23928 Homo sapiens serine hyd	gb_vt:1IHNGNJB	-	7.00	105.43	1.6e+03	1933	L40881 Infectious hematopoi
gb_vt:AA03190	-	7.00	106.77	1.4e+03	1602	AA03190 Asclaus asclaus mRNA	gb_vt:1IHNGNJB	-	7.00	105.41	1.6e+03	1933	L40882 Infectious hematopoi
gb_vt:BC018757	-	7.00	106.77	1.4e+03	1603	BC018757 Homo sapiens, clone M	gb_vt:AY062464	-	7.00	105.41	1.7e+03	1940	AY062464 Arabidopsis thalia
gb_vt:AX083337	-	7.00	106.75	1.4e+03	1606	AX083337 Sequence 29 from Pat	gb_vt:AF076619	-	7.00	105.37	1.7e+03	1950	AF076619 Rattus norvegicus
gb_vt:1IHNS0401	-	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_vt:AX036113	-	7.00	105.25	1.7e+03	1982	AX036113 Sequence 12 from P
gb_vt:1IHNSP	-	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_vt:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:1H015170	-	7.00	106.74	1.4e+03	1609	M1603 Infectious hematopoi	gb_vt:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat
gb_vt:AX151577	-	7.00	106.71	1.4e+03	1617	AX151577 Sequence 182 from Pat	gb_vt:158528	-	7.00	105.25	1.7e+03	1982	158528 Sequence 12 from pat

gb_ba:SCU60175	-	105.25	1.7e+03	1984	U60175 Sphingomonas chlorophen	gb_vi:SQOLVE	-	103.30	2.2e+03	2607	M38122 Squash leaf curl 11R
gb_ro:RNL4MLV1	-	105.24	1.7e+03	1986	U87602 Rattus norvegicus L1 re	gb_pat:AX036123	-	103.19	2.2e+03	2605	AX036123 Sequence 22 from p
gb_pat:RNL4MLV1	+	105.23	1.7e+03	1987	AX281642 Sequence 51 from Pat	gb_pat:158538	-	103.19	2.2e+03	2645	158538 Sequence 22 from pat
gb_ba:PEP0133737	+	105.19	1.7e+03	2000	AI133737 Paenibacillus polymy	gb_pat:187013	-	103.19	2.2e+03	2645	187013 Sequence 22 from pat
gb_pl:AY046011	-	105.16	1.7e+03	2009	AY046011 Arabidopsis thaliana	gb_ro:BC003668	-	103.16	2.2e+03	2658	BC003668 Homo sapiens, clone
gb_pr:BC022461	-	105.12	1.7e+03	2019	BC022461 Homo sapiens, neurog	gb_ro:BC021440	-	103.16	2.2e+03	2658	BC021440 Homo sapiens, clone
gb_pat:EE34346	+	105.10	1.7e+03	2026	E34346 DNA and process for pcc	gb_pr:AF276889	-	103.13	2.2e+03	2671	AF276889 Homo sapiens amaro
gb_in:AF216210	+	105.01	1.7e+03	2050	AF216210 Drosophila buzzatii	gb_pl:AB037203	-	103.12	2.2e+03	2671	AB037203 Glycyrhiza glabra
gb_pl:AY050818	+	105.00	1.7e+03	2054	AY050818 Arabidopsis thaliana	gb_ba:AF352332	-	103.11	2.2e+03	2676	AF352332 Candidatus Procoba
gb_ro:MUSC0142	+	105.00	1.7e+03	2054	D38412 Mouse mRNA for p10-1p	gb_ro:BC005672	-	103.11	2.2e+03	2677	BC005672 Mus musculus, clone
gb_pat:MUSC0142	+	104.98	1.7e+03	2060	AX036112 Sequence 11 from Pat	gb_ba:AF352333	-	103.10	2.2e+03	2678	AF352333 Candidatus Procoba
gb_pat:158527	-	104.98	1.7e+03	2060	158527 Sequence 11 from patent	gb_ov:IRU068757	+	103.10	2.2e+03	2680	IRU068757 Rana ridibundus secre
gb_pat:187002	-	104.98	1.7e+03	2060	187002 Sequence 11 from patent	gb_ov:XLSGII	+	103.06	2.2e+03	2696	XLSGII x laevis mRNA for se
gb_pl:AF418308	+	104.91	1.8e+03	2064	AF418308 Arabidopsis thaliana	gb_pr:AF038132	-	103.04	2.2e+03	2700	AF038132 Homo sapiens putat
gb_ro:BC013564	+	104.91	1.8e+03	2080	BC013564 Mus musculus, RIKEN	gb_in:IDMNR8	+	103.04	2.2e+03	2708	IDMNR8 E.melanogaster mRNA
gb_pl:SCNDJ033C	+	104.88	1.8e+03	2089	Z74081 S.cerevisiae chromosome	gb_ba:094413	-	103.03	2.2e+03	2708	094413 E.melanogaster mRNA
gb_pl:LES193712	+	104.79	1.8e+03	2115	AI133732 Arabidopsis thaliana	gb_in:AY040494	-	102.98	2.3e+03	2726	AY040494 Alveolinella quol
gb_pl:LES193712	+	104.77	1.8e+03	2122	AI133732 Arabidopsis thaliana	gb_ba:SPU6330	-	102.98	2.3e+03	2784	SPU6330 Synecococcus PCC700
gb_pr:HSAPNPK1	-	104.75	1.8e+03	2126	AI12910 Human Na ⁺ K ⁺ ATPase gen	gb_ba:RSEK0XG	-	102.79	2.3e+03	2800	RSEK0XG Rhizobium sp. exoA
gb_pl:AF306784	-	104.70	1.8e+03	2143	AF306784 Cochliobolus carbonum	gb_ba:BFU05415	-	102.74	2.3e+03	2819	BFU05415 Bacillus firmus puta
gb_pr:AF258591	-	104.69	1.8e+03	2145	AF258591 Homo sapiens p784 m	gb_pr:AF355121	-	102.72	2.3e+03	2836	AF355121 Pan troglodytes cl
gb_pat:EE52146	-	104.60	1.8e+03	2172	EE52146 Novel protein and DNA	gb_pl:AF467993	+	102.71	2.3e+03	2836	AF467993 Tolypocladium Infl
gb_pat:EE52150	-	104.60	1.8e+03	2172	EE52150 Novel protein and DNA	gb_pr:AF355118	+	102.70	2.3e+03	2835	AF355118 Pan troglodytes cl
gb_pr:AF270549	+	104.53	1.8e+03	2193	AF270549 Homo sapiens clone 18	gb_pat:AF355126	+	102.69	2.3e+03	2838	AF355126 Homo sapiens clone
gb_pr:HSOC0145	+	104.51	1.9e+03	2199	U41068 Human collagen alpha2(X	gb_pr:AF355117	-	102.67	2.3e+03	2845	AF355117 Pan troglodytes cl
gb_pat:AX317850	+	104.51	1.9e+03	2200	AX317850 Sequence 25 from Pat	gb_pr:AF355117	-	102.67	2.3e+03	2845	AF355117 Pan troglodytes cl
gb_pat:AX353714	+	104.51	1.9e+03	2204	AX353714 Sequence 7 from Pat	gb_pr:AF355120	-	102.67	2.3e+03	2847	AF355120 Pan troglodytes cl
gb_pr:AB047835	-	104.46	1.9e+03	2215	AB047835 Macaca fascicularis	gb_pr:AB072792	+	102.65	2.4e+03	2853	AB072792 Macaca fasciculari
gb_ba:SYO514	-	104.44	1.9e+03	2227	M82953 Synecococcus sp. IslA	gb_pr:AB072792	+	102.65	2.4e+03	2853	AB072792 Macaca fasciculari
gb_pr:HUMMEL18	+	104.42	1.9e+03	2230	U41068 Human collagen alpha2(X	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ba:VIBEMET8	+	104.41	1.9e+03	2230	X68658 Vibrio sp. gene for V	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pl:VEVCINMTR	-	104.37	1.9e+03	2242	X49831 V.faba VEVGINV mRNA for	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_com:OCU8193	+	104.29	1.9e+03	2267	U38193 Oryctolagus cuniculus	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ro:AF024630	+	104.23	1.9e+03	2286	AF024630 Mus musculus gamma-am	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ro:RNL4MLV1	-	104.23	1.9e+03	2288	X06942 Rat A-raf mRNA, 9/1/93	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ba:BAOCRI10	-	104.23	1.9e+03	2288	X07687 Rat L1R B7 repetitive	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ba:BAOCRI10	-	104.22	1.9e+03	2290	M04475 B.thuringiensis replicat	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pl:AF115279	-	104.22	1.9e+03	2290	AF115279 Euglena agilis small	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:AF448510	-	104.06	2.0e+03	2344	AF448510 Homo sapiens histone	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_in:IBR239060	+	104.05	2.0e+03	2344	AF332141 Mus musculus receptor	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pl:AB023960	+	103.97	2.0e+03	2373	AB023960 Trypanosoma brucei B5	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:HUMR814R	+	103.96	2.0e+03	2376	L76687 Homo sapiens Gb14 mRNA	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:AB063576	-	103.96	2.0e+03	2376	AB063576 Homo sapiens Gb14 mRNA	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:AB063576	-	103.92	2.0e+03	2389	AB063576 Homo sapiens Gb14 mRNA	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ba:AF092086	+	103.88	2.0e+03	2400	AF092086 Bacillus licheniform	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_in:AF092086	+	103.88	2.0e+03	2404	AF092086 Drosophila melanogast	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ov:AF098788	+	103.87	2.0e+03	2404	AF098788 Gallus gallus nucleat	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ba:AF079316	+	103.87	2.0e+03	2407	AF079316 Shigella flexneri str	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:BC007979	+	103.84	2.0e+03	2416	BC007979 Homo sapiens, serine	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ro:BC021840	+	103.84	2.0e+03	2416	BC021840 Mus musculus, clone	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ro:AF428262	+	103.81	2.0e+03	2425	AF428262 Mus musculus calcitri	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:AK023141	-	103.81	2.0e+03	2427	AK023141 Homo sapiens CDNA FL	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ba:BMGLE	+	103.64	2.1e+03	2484	X37448 B.mellensis gale, mazg	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pat:EE52145	+	103.60	2.1e+03	2499	EE52145 Novel protein and DNA	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pat:EE52149	+	103.60	2.1e+03	2500	EE52149 Novel protein and DNA	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_in:AF149286	+	103.59	2.1e+03	2503	AF149286 Caenorhabditis elegat	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:AK001615	+	103.58	2.1e+03	2506	AK001615 Homo sapiens CDNA FL	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pl:PSA0M1P1G	+	103.58	2.1e+03	2506	AF223312 Plinus sylvestris mld	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pat:AX149316	+	103.56	2.1e+03	2511	AX149316 Tomato (L.esculentum)	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_hlg:AC098341	+	103.55	2.1e+03	2515	AC098341 Sequence 1 from Pat	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pl:LES03343	+	103.54	2.1e+03	2519	LES03343 Lycopersicon esculent	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:HSK000385	+	103.52	2.1e+03	2521	AL080084 Homo sapiens mRNA; c	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ba:STIMPABC	+	103.52	2.1e+03	2527	X53528 Salmonella typhimurium	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pl:AY0455919	-	103.48	2.1e+03	2527	AY0455919 Arabidopsis thaliana	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:AF162667	-	103.46	2.1e+03	2547	AF162667 Homo sapiens mRNA for	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ov:AB019790	-	103.46	2.1e+03	2547	AB019790 Gallus gallus mRNA for	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ba:AF227554	-	103.44	2.1e+03	2547	AF227554 Porphyromonas gingiv	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:AK056588	-	103.44	2.1e+03	2556	AK056588 Homo sapiens CDNA FL	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_ov:1C5STS8P	-	103.37	2.1e+03	2580	X16078 T. californica mRNA for	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for
gb_pr:BC018109	-	103.36	2.1e+03	2585	BC018109 Homo sapiens, Stimlat	gb_ba:BP3C35	-	102.63	2.4e+03	2864	BP3C35 B.pertusis gene for


```

gb_pl:AF053127 - 7.00 101.39 2.8e+03 3405 | AF053127 Malus domestica leucl
gb_to:AF334942 - 7.00 101.30 2.8e+03 3449 | AF334942 Mus musculus toll-11k
gb_to:AF334943 - 7.00 101.27 2.8e+03 3463 | AF334943 Mus musculus toll-11k
gb_to:AF334944 - 7.00 101.23 2.8e+03 3483 | AF334944 Mus musculus toll-11k
gb_pat:AR139174 + 7.00 101.22 2.8e+03 3487 | AR139174 Sequence 32 from pat
gb_to:MMK00BRB - 7.00 101.20 2.8e+03 3499 | Y10296 Mus musculus mRNA for 1
gb_to:MMN00BRB - 7.00 101.20 2.8e+03 3499 | Y10296 Mus musculus mRNA for 1
gb_to:AF245455 - 7.00 101.18 2.8e+03 3507 | AF245455 Drosophila melanogast
gb_to:MG249993 - 7.00 101.13 2.9e+03 3530 | AJ249993 Mytilus gallioprovinci
gb_to:AV033326 - 7.00 101.10 2.9e+03 3546 | AV033326 Uncultured gamma prot
gb_pl:SCYAK1 - 7.00 101.00 2.9e+03 3599 | Y10056 Yeast YAK1 gene for yak
gb_to:AF291054 + 7.00 100.96 2.9e+03 3617 | AF291054 Mus musculus 8-oxodg
seq_name: gb_ba:AF242417

seq_documentation_block:
LOCUS AF242417 777 bp DNA linear BCT 01-JAN-2002
DEFINITION Escherichia coli heat-labile enterotoxin subunit A precursor, gene,
complete cds.
ACCESSION AF242417 GI:18026885
VERSION AF242417.1 GI:18026885
KEYWORDS
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 777)
AUTHORS Chang G.N. and Ho K.C.
TITLE Heat-labile enterotoxin subunit A gene of Escherichia coli
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Chang G.N. and Ho K.C.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2000) Botany, National Taiwan University 1
JOURNAL Roosevelt Road, Sec. 4, Taipei, Taiwan 106, Republic of China
FEATURES
Source
1. 777
/organism="Escherichia coli"
/strain="CH-C2"
/serotype="O-136"
/specific_host="pig"
/db_xref="taxon:562"
/country="Taiwan"
/note="Isolated from hemolytic cells"
1. 777
/note="ltpa"
/codon_start=1
/transl_table=1
/product="heat-labile enterotoxin subunit A precursor"
/protein_id="AAL55671.1"
/db_xref="GI:18026886"
/transl_table="MKMTFFIFLLASPLRYANGDRLYRADSRPPEIKRSGIMPRG
HNEYFDRCOMNINLYDHARGTGTGTPYRVDYVSTSLRSALHAGOSLSTSTY
IYIATAPNENYNDVIGVSPHYEDVSALGIPYSQITGMTRVFGVIDELHFN
REYDRYRLNLNAPADGVRLAGPPDQAMKEEPIHAPGCGSSSTITIGDYCN
EETONLSTIYLRVQSKVKRQIFSDYSEVDIYNRIKDEL"
1. 54
sig_peptide
55. 630
mat_peptide
/product="heat-labile enterotoxin subunit A1"
637. 774
mat_peptide
/product="heat-labile enterotoxin subunit A2"
BASE COUNT 255 a 136 c 164 g 222 t
ORIGIN

```

```

alignment_scores:
Quality: 240.00 Length: 240
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x AF242417

```

```

Align seg 1/1 to: AF242417 from: 1 to: 777
1 ASNGIYASPAAGLEUTYRARGALASPSERARPRPROASPGIULELY 17
|||||
55 AATGCGCAGATATATACCGTGTACTGACCTAGACCCAGATGAAATPAA 104
|||||
17 SARSGERGLYLEUWETPROARGLYHISASNGIUTYRPHASPARG 34
|||||
105 ACGTTCGCGAGGTGTATATGCGCAGGCGCATATAGACTTCGATGAG 154
|||||
34 IYTHGLMETASNLLEASNLEUTYRASPHISALATRGLYTHGLNTH 50
|||||
155 GAACCAATGATATATATCTTATGATCAGCGGAGGAGCAACAAACC 204
|||||
51 GLYPHEVALARGTYRASPSGLYTYRVALSERTHSRLEUSERLEAR 67
|||||
205 GCGTTTGTACATATATGATGATGATGATGATGATGATGATGATGAT 254
|||||
67 GSERALHISLEUALAGLYNSERILEUSERCGLYTYRSETHRYT 84
|||||
255 AAGTCTCAGCTTAGCAGACAGCTCATATATATGATGATGATGATGAT 304
|||||
84 YRLETYRVALILEALATRALAPROASNMETPHEASNVALASNPVAL 100
|||||
305 ATATATATGATTATAGCAGACACCAATATGTTAATGTTAATGATCTA 354
|||||
101 LEUGLYVALTYRSEPRHISPROTYRGLINGIUVASERALEUGI 117
|||||
355 TTAGGCGTATACGCCCTCCACCATGATGACAGGAGGTTTGGCTTAGG 404
|||||
117 YGLYLEPROTYRSEGINILETYRGLYTYRPTYRARGVALASNPHEGLY 134
|||||
405 TGGATATACCATATCTCATATATATGATGATGATGATGATGATGATG 454
|||||
134 ALLIENSPGLNARGLEUHIISARGASNARGIUTYRARGASPARGYTYR 150
|||||
455 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 504
|||||
151 ARGASNLEUASNILEALPROALAGLUSPGIYTYRARGLEUALAGLYPH 167
|||||
505 AGAATATGATATATATGCTCCGCGCAGGAGATGTTACATATACGAGTT 554
|||||
167 EPRPROASPHISGINALATYRARGGLUGLUPTOTRPLEHISHISALAP 184
|||||
555 CCCACCGGATCCACCAAGCTTGAGAGAGAACCCGATTCATCATCGAC 604
|||||
184 TGGIINGLYCYGLYASNSERSEARATHRIEMHGLYASPTHCYASNS 200
|||||
605 CACAAGGTTGTGGAAATTCATCAAGAACCAATCAGAGTGATGATGAT 654
|||||
201 GIUGIUTRGINASNLSESTHRIETYLEUARGGLUYTYRGINSESLY 217
|||||
655 GAGGAGACCCAGATCTAGACATATATATCTCAGGAGATTCATCAATA 704
|||||
217 SVALLYSARGINILEPHESEASPTYRGLINSEGLUVALASPLETYRA 234
|||||
705 AGTTAAGAGCGCATATTTTCAGACTATCAGAGGTGACATATATATA 754
|||||
234 SNARGILEARGLASPSGLU 240
|||||
755 ACAGAAATTCGGGATGATTA 774
|||||
seq_name: gb_ba:P307LTA

seq_documentation_block:
LOCUS P307LTA 777 bp DNA linear BCT 24-APR-1996
DEFINITION Plasmid P307 (from E. coli) heat-labile enterotoxin subunit A (lta)
gene, complete cds.
ACCESSION M35581
VERSION M35581.1 GI:150458
KEYWORDS enterotoxin; lta gene.
SOURCE Plasmid P307 (clone: PAT153.) DNA.

```


ORGANISM Plasmid P307
 plasmids.
 1 (bases 1 to 777)
 REFERENCE Dykes,C.W., Halliday,I.J., Hobden,A.N., Read,M.J. and Harford,S.
 AUTHORS A comparison of the nucleotide sequence of the A subunit of heat-
 TITLE labile enterotoxin and Cholera toxin
 JOURNAL FEMS Microbiol. Lett. 26, 171-174 (1985)
 FEATURES
 source
 1..777
 /organism="Plasmid P307"
 /plasmid="Plasmid P307"
 /specific_host="Escherichia coli"
 /db_xref="taxon:2472"
 /clone="PAT153."
 1..777
 /gene="lta"
 1..777
 /gene="lta"
 CDS
 1..777
 /gene="lta"
 /codon_start=1
 /transl_table=1
 /product="heat-labile enterotoxin subunit A"
 /protein_id="AA98202.1"
 /db_xref="GI:150459"
 /translation="MKNTFFIFILLASPLYANGDRLYRADSRPPDEIKRSGGLMPRG
 HNEYDRGTONMINLYDHARGTQTFVARDGYSVSTLSRAHLAGOSILSGYSTY
 IYVATAPNMENVNDVLGVSPHPYEOEVSALGIPYQIIGWYRVNFGVIDERLARN
 REYDRYRNLIAPAEDEGYRLAGPRPDQAMREPMIHHAPQGGCNSRITTDTCN
 EETONLSTIYREYOSKVKROIYFSQYQSEVDIYNIRDEL"
 BASE COUNT 255 a 136 c 164 g 222 t
 ORIGIN
 alignment_scores:
 Quality: 240.00 Length: 240
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-528-682-1 x P307LTA ..
 Align seg 1/1 to: P307LTA from: 1 to: 777
 1 ASNGIYAspargLeuTYrArGAlAspSerArGProPAspGluIleY 17
 |||||
 55 AATGGCGACAGATTATACCGTCTGACTAGACCCCAATGAAATAA 104
 |||||
 17 sArGserGIyGlyLeuMePrOArGlyHIsaNGIuTYrPhesArpArG 34
 |||||
 105 ACGTTCGGAGGCTTATGCCAGAGGCAATGAGTACTTCGATGAG 154
 |||||
 34 lYthrgInMeLAsnIleAsnLeuTYrAspHIsAarGlythrgInLthr 50
 |||||
 155 GAACCTCAATGAATTAATCTTATGATCAGCGAGAGAACCAAC 204
 |||||
 51 GIYPhEValArGTyrAspAspGIYTYrValSerThrSerLeuSerLeuAr 67
 |||||
 205 GCGTTTGTCAATATGATGACGATATGTTCCACTTCTTATGTTGAG 254
 |||||
 67 gSerAlaHIsLeuAlaGIyInSerIleLeuSerGIYTYrSerThrTYrT 84
 |||||
 255 AAGTTCACACTTAGCAGACAGCTTATATATCAGATATTCCTACTACT 304
 |||||
 84 YtlEtyrValIleAlaThraLProAsnMetPhEasnValAspVal 100
 |||||
 305 ATATATATGTTATAGCAGACACCAATATGTTTAAATGTTAAGATGA 354
 |||||
 101 LeuGIYValTYrSerProHIsProTYrGIuNGIuValSerAlaLeuGI 117
 |||||
 355 TTAGGCGGTATACAGCCCTCACCATATGACAGAGAGTTCTCGCTTAGG 404
 |||||
 117 YGIYlEPrOtyrSerGIuNIleTYrGIYTYrArGValAsnPhEGIY 134
 |||||
 405 TCGAATACCATATTCAGATATATGATGATCGTATCGTTAAATTTGCTG 454

134 AlIleAspGIuArGLeuHIsArGAsnArGLeuIuTYrArGAspArGTYrTYr 150
 |||||
 455 TGATGATGAAACGATTATCATCGTACAGGCAATATATAGACCGGATATAC 504
 |||||
 151 ArGAsnLeuAsnIleAlaPrOAlaGIuAspGIYTYrArGLeuAlaGIY 167
 |||||
 505 AGAATGTGAATATAGCTCCGGCAGAGAGATGTTACGATTTACAGAGTTT 554
 |||||
 167 ePrOPrOAspHIsGIuAlaTPArGGIuGIuPrOTPrIleHIsAlaP 184
 |||||
 555 CCCACCCGATCACCAGCTTGAGAGAAACCTGGATTCAATCAGCAGAC 604
 |||||
 184 rOGInGIYcysGIYAsnSerSerArGThrIlethrgInAspPthrcysAsn 200
 |||||
 605 CACAGAGTTGTGAAATTCATCAAGAACATCACAGGTGATCTTATAT 654
 |||||
 201 GIuGIuThrgInAsnLeuSerThrIleTYrLeuArGLeuIuTYrGInSerly 217
 |||||
 655 GAGAGAGCCAGAAATCGACACATATATCTCAGGCAATATCATCAACAA 704
 |||||
 217 sValYsArGIuNIlePheserAspTYrGInSerGIuValAspIleTYrA 234
 |||||
 705 AGTTAAGAGCGACATATTTTCAGACTATCATGACAGAGTTGACATATATA 754
 |||||
 234 sNArGIleArGAspGIuLeu 240
 |||||
 755 ACAGAAATTCGGAGATGAATTA 774
 seq_name: gb_pat:A04913
 seq_documentation_block:
 LOCUS A04913 777 bp DNA linear PAT 13-JUL-1993
 DEFINITION E.coli LTA gene for labile toxin A.
 ACCESSION A04913
 VERSION A04913.1 GI:412215
 KEYWORDS labile toxin A.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1 to 777)
 AUTHORS Hayes,M.V., Harford,S. and Ross,G.W.
 TITLE Microbiological process
 JOURNAL Patent: EP 0145486-A 5 19-JUN-1985;
 GLAXO GROUP LIMITED
 FEATURES
 source
 1..777
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1..777
 /codon_start=1
 /transl_table=1
 /product="labile toxin A"
 /protein_id="CAA00402.1"
 /db_xref="GI:412216"
 /db_xref="SWISS-PROT:P06717"
 /translation="MKNTFFIFILLASPLYANGDRLYRADSRPPDEIKRSGGLMPRG
 HNEYDRGTONMINLYDHARGTQTFVARDGYSVSTLSRAHLAGOSILSGYSTY
 IYVATAPNMENVNDVLGVSPHPYEOEVSALGIPYQIIGWYRVNFGVIDERLARN
 REYDRYRNLIAPAEDEGYRLAGPRPDQAMREPMIHHAPQGGCNSRITTDTCN
 EETONLSTIYREYOSKVKROIYFSQYQSEVDIYNIRDEL"
 BASE COUNT 255 a 136 c 164 g 222 t
 ORIGIN
 alignment_scores:
 Quality: 240.00 Length: 240
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-528-682-1 x A04913 ..

Align seg 1/1 to: A04913 from: 1 to: 777

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
|||||
55 AATGGCGACAGATATATACCGTGCACCTTGACCCCGAGATGAATAA 104
|||||
17 SarSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
105 ACGTTCCGAGGTCTTATGCGGAGGCGCATATGAGTACTTCGATAG 154
|||||
34 LyrThrGluMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
155 GAACCAATGAATATATATCTTATGATACGCGAGAGAACACAAACC 204
|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
|||||
205 GCGTTTGTCAGATATGATGACGAGATGTTTCCACTTCTTAGTTGAG 254
|||||
67 gSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyrSerThrTyrT 84
|||||
255 AAGTCCCTACTTAGCAGACAGTCTATATTCAGGATATCCACTTACT 304
|||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
305 ATATATATGTTATAGCGACACACCAATATGTTAAATGTTAATGATGTA 354
|||||
101 LeuGlyValIleTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||
355 TTAGGCGGTATACAGCCCTCACCCATATGACAGAGGTTTCTGCGTTAG 404
|||||
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
|||||
405 TGGATATACCATATCTCAGATATATGATGATGATGATGATGATGATG 454
|||||
134 AlIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspArgTyrTyr 150
|||||
455 TGATTGATGAACGATTACATCTGTAACAGGAATATAGACCGGATATTAC 504
|||||
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||
505 AGAATATCGAATATATAGCTCCGCGACAGAGATGTTACATATAGCAGGTT 554
|||||
167 eProProAspHisGlnAlaTyrPargGluGluProTyrIleHisAlaP 184
|||||
555 CCCACCGGATCACCAAGCTTGAGAGAAACCCGATCATCATGAC 604
|||||
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
|||||
605 CACAGGTTGTGGAAATTCATCAAGAACACACAGGTGATCTTGTAAT 654
|||||
201 GluGluTyrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSerI 217
|||||
655 GAGGAGACCCAGAAATCTGACACAAATATCTCAGGAAATATCAATCAA 704
|||||
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
|||||
705 AGTTAAGAGCGAGATATTTTCAGACTATCAGTCAAGAGGTTGACATATTA 754
|||||
234 snArgIleArgAspGluLeu 240
|||||
755 ACAGATTCGGGATGAATTA 774

```

seq_name: gb_ba:ENSLTPA

seq documentation block:

LOCUS ENSLTPA 866 bp DNA linear BCT 26-Apr-1993
 DEFINITION Plasmid ENT P307 (enterotoxigenic E.coli) LTP subunit A gene.
 ACCESSION M15362
 VERSION M15362.1 GI:148337
 KEYWORDS LTP gene; heat-labile toxin.
 SOURCE Plasmid ENT (enterotoxigenic E.coli porcine isolate) DNA, clone p370.

ORGANISM Plasmid ENT
 plasmids.
 REFERENCE 1 (bases 1 to 866)
 AUTHORS Yamamoto, T., Gojibori, T. and Yokota, T.
 TITLE Evolutionary origin of pathogenic determinants in enterotoxigenic
 Escherichia coli and Vibrio cholerae O1
 JOURNAL J. Bacteriol. 169, 1352-1357 (1987)
 MEDLINE 87137303
 COMMENT Clean copy of sequence [1] kindly provided by T.Yamamoto
 (22-Apr-1987)

FEATURES

source location/Qualifiers
 1..866 /organism="Plasmid ENT"
 /plasmid="plasmid ENT"
 /db_xref="taxon:2463"
 90..143
 sig_peptide
 /note="Heat-labile toxin subunit A signal peptide"
 90..866
 CDS
 /note="Heat-labile toxin subunit A precursor"
 /codon_start=1
 /protein_id="AAA24793.1"
 /transl_table=11
 /db_xref="GI:148338"

mat_peptide
 /product="Heat-labile toxin subunit A"
 277 a 149 c 181 g 259 t
 BASE COUNT 277 a 149 c 181 g 259 t
 ORIGIN unreported.

alignment_scores:
 Quality: 240.00 Length: 240
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ENSLTPA ..

Align seg 1/1 to: ENSLTPA from: 1 to: 866

```

1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProProAspGluIleLeu 17
|||||
144 AATGGCGACAGATATATACCGTGCACCTTGACCCCGAGATGAATAA 193
|||||
17 SarSerGlyGlyLeuMetProArgGlyHisAsnGluTyrPheAspArg 34
|||||
194 ACGTTCCGAGGTCTTATGCGGAGGCGCATATGAGTACTTCGATAGAG 243
|||||
34 LyrThrGluMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
|||||
244 GAACCTCAATGAATATATATCTTATGATCAGCGAGAGAGAACCAAAC 293
|||||
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeuArg 67
|||||
294 GCGTTTGTCAGATATGATGACGAGATGTTTCCACTTCTTAGTTGAG 343
|||||
67 gSerAlaHisLeuAlaGlyInserIleLeuSerGlyTyrSerThrTyrT 84
|||||
344 AAGTCCCTACTTAGCAGACAGTCTATATATGATGATGATGATGATG 393
|||||
84 YrIleTyrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
394 ATATATATGTTATAGCGACACCAATATGTTAAATGTTAATGATGTA 443
|||||
101 LeuGlyValIleTyrSerProHisProTyrGluGlnGluValSerAlaLeuG 117
|||||
444 TTAGGCGGTATACAGCCCTCACCCATATGACAGAGGTTTCTGCGTTAG 493
|||||
117 YGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134

```


|||||
494 TGAATACCATTTCTCAGATATATGATGCTATCCGTTAATTTGCTG 543
134 a11leaspgluargleuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
|||||
544 TGATTGATGACGATTACATCTGTAACAGGGAATATAGACACCGTATTAC 593
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
|||||
594 AGAATCTGAATATAGCTCCGCGAGAGATGTTACGATTTACAGAGT 643
167 eProPAspHisGlnAlaIleTyrArgGluGluProTyrIleHisIleAla 184
644 CCCACCGGATCACCAAGCTTGAGAGAGAACCCTGGATTCAATCATGAC 693
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
694 CACAGAGTTGTGAAATTCATCAAGAACATCACAGCTGATCTGTAA 743
201 G1uG1uThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
|||||
744 GAGGAGACCCAGATCTGACACAAATATCTCAGGGAATATCAATCAA 793
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
794 AGTTAAGAGCAGATATTTTCAGACTATCATGTCAGAGCTTGACATATA 843
234 snArgIleArgAspGluLeu 240
|||||
844 ACAGAAATTCGGATGAATTA 863
seq_name: gb_ba:ENRLTPA
seq_documentation_block:
LOCUS ENRLTPA 898 bp DNA linear BCT 26-APR-1993
DEFINITION Plasmid ENM-R PCG86 (enterotoxigenic E.coli) LTP subunit A gene.
ACCESSION M15361
VERSION M15361.1 GI:148333
KEYWORDS LTP gene; heat-labile toxin.
SOURCE Plasmid ENM-R (enterotoxigenic E.coli porcine isolate) DNA, clone PCG86.
ORGANISM Plasmid ENM-R
REFERENCE 1 (bases 1 to 898)
AUTHORS Yamamoto,T., Gotohori,T. and Yokota,T.
TITLE Evolutionary origin of pathogenic determinants in enterotoxigenic Escherichia coli and Vibrio cholerae O1
JOURNAL J. Bacteriol. 169, 1352-1357 (1987)
MEDLINE 87137303
COMMENT Clean copy of sequence [1] kindly provided by T.Yamamoto (22-APR-1987).
FEATURES
source Location/Qualifiers
1..898
/organism="Plasmid ENM-R"
/plasmid="Plasmid ENM-R"
/db_xref="taxon:2462"
sig_peptide 122..175
122..175
/note="Heat labile toxin subunit A signal peptide"
CDS 122..898
122..898
/note="Heat labile toxin subunit A precursor"
/codon_start=1
/transl_table=11
/protein_id="AAA24791.1"
/db_xref="GI:148334"
/translation="MKNTFFIFILLASPLYANGDRLYRADSPPDEIKRSGGLPRG HNEFYDCTQNNINLYDHARGTQTFVRYDGYSTSLSAHLACGSIISGYSTY IYVATAPNMFVNVDLVGVSPHPEQGVSDVSGIYSGIYGVNFGVIDERLHN REYRDYRNINIAPEADGVYRAGPPDHOAMREPMWIIHHAPOGGGNSRRTTGTCTN EETQNLSTIYLRVQSKVRQIFSDYQSEVDIYNRIREL"
mat_peptide 176..895
176..895
/product="Heat labile toxin subunit A"
BASE COUNT 289 a 155 c 186 g 268 t
ORIGIN Unreported.

alignment_scores:
Quality: 240.00 Length: 240
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x ENRLTPA ..
Align seq 1/1 to: ENRLTPA from: 1 to: 898
1 AsnGlyAspArgLeuTyrArgAlaAspSerArgProPAspGluIle 17
|||||
176 AATGGCGACAGATTATACCTGCTGACTTACAGCCCAATGAAATAA 225
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyTyrPheAspArg 34
226 ACGTCCGAGAGTCTTATGCCAGAGGCAATATGACTTCTGATAGAG 275
34 IythrGlnMetAsnIleAsnLeuTyrAspHisAlaArgGlyThrGlnThr 50
276 GAACCTCAATGAATATTAATCTTTATGATCAGCGAGAGAACCAACC 325
51 GlyPheValArgTyrAspAspGlyTyrValSerThrSerLeuSerLeu 67
326 GCCTTTCTCAGATATGATGACGGATATGTTCCACTTCTTACTGATGAG 375
67 gSerAlaHisLeuValArgIleSerIleLeuSerGlyTyrSerThrTyr 84
376 AAGTCTCACTTAGCAGAGACAGCTATATATGAGATATCCACTACT 425
84 YrlleTyrValIleAlaThrAlaProAsnMetPheAsnValAspVal 100
426 ATATATATGTTATAGCGACACCAATATGTTTAAATGTAATGATGA 475
101 LeuGlyValIleTyrSerProHisProTyrGluGlnGluValSerAlaLeu 117
476 TTAGGCTATACAGCCCTCACCCATATGAACAGAGGTTTCTGGTTAGG 525
117 yGlyIleProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPheGly 134
526 TGAATATCAATATCTCAGATATGATGATGATGATGATGATGATGATG 575
134 a11leaspgluargleuHisArgAsnArgGluTyrArgAspArgTyrTyr 150
576 TGATTGATGAACGATTACATCTGTAACAGGGAATATAGACCCGTTATTAC 625
151 ArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPh 167
626 AGAATCTGAATATAGCTCCGCGAGAGATGTTACGATTTAGCAGGTTT 675
167 eProPAspHisGlnAlaIleTyrArgGluGluProTyrIleHisIleAla 184
676 CCCACCGGATCACCAAGCTTGAGAGAGAACCCTGGATTCAATCATGAC 725
184 roGlnGlyCysGlyAsnSerSerArgThrIleThrGlyAspThrCysAsn 200
726 CACAGAGTTGTGAAATTCATCAAGAACATCACAGCTGATCTGTAA 775
201 G1uG1uThrGlnAsnLeuSerThrIleTyrLeuArgGluTyrGlnSer 217
776 GAGGAGACCCAGATCTGACACAAATATCTCAGGGAATATCAATCAA 825
217 sValIysArgGlnIlePheSerAspTyrGlnSerGluValAspIleTyr 234
826 AGTTAAGAGCAGATATTTTCAGACTATCATGTCAGAGCTTGACATATA 875
234 snArgIleArgAspGluLeu 240
|||||
876 ACAGAAATTCGGATGAATTA 895
seq_name: gb_pat:E03421


```

seq_documentation_block:
  LOCUS      E03421              1148 bp    DNA          linear    PAT 29-SEP-1997
  DEFINITION  DNA encoding Lth of entero toxigenic Escherichia coli.
  ACCESSION  E03421
  VERSION    E03421.1 GI:2171637
  KEYWORDS   JP 1992079898-A/1.
  SOURCE      Escherichia coli.
  ORGANISM    Escherichia coli
               Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
               Escherichia.
REFERENCE    1 (bases 1 to 1148)
  AUTHORS    Danbara,H. and Abe,A.
  TITLE      DNA AND RNA PROBE TO SIMULTANEOUSLY DETECT VIBRIO CHOLERAE AND
              ENTEROTOXIGENIC ESCHERICHIA COLI AND DETECTION OF VIBRIO CHOLERA
              AND ENTEROTOXIGENIC ESCHERICHIA COLI USING THE SAME DNA AND RNA
              PROBE
  JOURNAL    Patent: JP 1992079898-A 1 13-MAR-1992;
              KITASATO INST.:THE
              OS      Escherichia coli
              PN      JP 1992079898-A/1
              PD      13-MAR-1992
              PF      23-JUL-1990 JP 1990194208
              PI      DANBARA HIROFUMI, ABE AKIO
              PC      C1201/68,C07H21/02,C07H21/04,C12N15/11,C12N15/31,C1201/04, PC
              PC      C1201/10,
              PC      (C12N15/31,C12R1:19),(C1201/04,C12R1:63,C12R1:19); CC
              CC      strandedness: Double;
              CC      topology: Linear;
              CC      hypothetical: No;
              CC      anti-sense: No;
              CC      *source: clone-PKAD008;
              FH      Key      Location/Qualifiers
              FT      misc_feature      1..1148
                                   /note="DNA fragment encoding Lth of entero
                                   /taxid=1148
                                   /organism="Escherichia coli"
                                   /db_xref="taxon:562"
              FT      sig_peptide      1..54
                                   /product="subunit A of Lth"
              FT      mat_peptide      55..774
                                   /product="subunit A of Lth"
              FT      CDS              1..777
                                   /product="precursor subunit A of Lth"
              FT      sig_peptide      774..837
                                   /product="subunit B of Lth"
              FT      mat_peptide      838..1145
                                   /product="subunit B of Lth"
              FT      CDS              774..1148
                                   /product="precursor of subunit B of Lth"
  FEATURES
    source      1..1148
                 Location/Qualifiers
    BASE COUNT  399 a      205 c      229 g      315 t
    ORIGIN
alignment_scores:
  Quality: 208.00      Length: 208
  Ratio: 1.000          Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
  US-09-528-682-1 x E03421 ..
Align seg 1/1 to: E03421 from: 1 to: 1148
5 LeuTyrAlaGlaaSPSerArgProProAspGluIleLysArgSerGlyG1 21
|||||
67 TTATACCGTGTGACTAGACCCCGAGATGAAATAAACGTCGCGAGG 116
|||||
21 YleuMetProArgGlyHisAsnGluTyrPheAspArgGlyThrGlnMet 38
|||||
117 TCTTATGCCAGAGGCGATTAATGAGTCTTCGATGAGAGAACTCAATGA 166
|||||

```

```

38 snlleasleuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
167 ATATTATCTTTATGATCAGCGAGAGAGACAAACCGGCTTTGTCAGA 216
|||||
55 TyrAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisI 71
|||||
217 TATGATGACGGATATGTTTCACCTCTTCTTCTTCTTCTTCTTCTTCT 266
|||||
71 uAlaGlyGlnSerLeuSerGlyTyrSerThrThrTyrTyrIleTyrVal 88
|||||
267 AGCAGACAGCTATATATATGAGATATTCACCTTACTATATATATGTA 316
|||||
88 leaAlaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTyr 104
|||||
317 TAGCGACAGACCAATATGTTTATGTTATGTTATGTTATGTTATGTTAT 366
|||||
105 SerProHisProTyrGluGluGluValSerAlaLeuGlyGlyLeuPro 121
|||||
367 AGCCCTCACCACCATATGAAACAGAGAGGTTTCGCGTTAGGTGGAATACCAT 416
|||||
121 rSerGlnIleTyrGlyTyrPheTyrAlaAsnPheGlyValIleAspGlu 138
|||||
417 TTTCAGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 466
|||||
138 rGleuHisArgAsnArgGlyTyrArgAspArgTyrTyrArgAsnLeuAsn 154
|||||
467 GATTACATCGTAAACAGAGGATATAGAGACCGGATATACAGAAATGCGAAT 516
|||||
155 lLeaAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAsp 171
|||||
517 ATAGCTTCGGCAGAGAGAGATGTTACAGATGTTACAGAGTTCCACCGGATCA 566
|||||
171 sGlnAlaThrArgGluGluProTrpPheHisAlaProGlnGlyCysG 188
|||||
567 CCAAGCTTGAGAGACAGACACCTGATTCATCATGACCAAGGTTGTC 616
|||||
188 lYasSerSerArgThrIleThrGlyAspPheCysAsnGluThrGln 204
|||||
617 GAATTCATCATGAGACAAATTAACAGGTGATCTGTATATGAGAGACCCAG 666
|||||
205 AsnLeuSerThrIleTyrLeuArg 212
|||||
667 AATCTGAGCACAAATATATCTCAGG 690
|||||
seq_name: gb_ba:AB011677
seq_documentation_block:
  LOCUS      AB011677              1434 bp    DNA          linear    BCT 23-APR-1998
  DEFINITION  Escherichia coli genes for heat-labile enterotoxin A subunit and B
              subunit, complete cds.
  ACCESSION  AB011677
  VERSION    AB011677.1 GI:3062900
  KEYWORDS   heat-labile enterotoxin B subunit; Lth b subunit; heat-labile
              enterotoxin A subunit; Lth a subunit.
  SOURCE      Escherichia coli (strain:1032 (Enterotoxigenic)) DNA.
  ORGANISM    Escherichia coli
               Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
               Escherichia.
REFERENCE    1 (bases 1 to 1434)
  AUTHORS    Komase,K.
  TITLE      Direct Submission
  JOURNAL    Submitted (02-MAR-1998) Katsuhiko Komase, The Kitasato Institute,
              Center for Basic Research, Lab. of Virology, Shirokane 5-9-1,
              Minato-ku, Tokyo 108-8642, Japan (E-mail:komase-k@kitasato.or.jp,
              Tel:+81-3-3444-6161, Fax:+81-3-3444-6637)
  2 (sites)
  Tamura,S., Asanuma,H., Tomita,T., Komase,K., Kawahara,K.,
  Danbara,H., Hattori,N., Watanabe,K., Suzuki,Y., Nagamine,T.,
  Aizawa,C., Oya,A. and Kurata,T.
  Escherichia coli heat-labile enterotoxin B subunits supplemented
  with a trace amount of the holotoxin as an adjuvant for nasal
  influenza vaccine

```


JOURNAL Vaccine 12 (12), 1083-1089 (1994)
 MEDLINE 95091056
 FEATURES
 source Location/Qualifiers
 1.1434
 /organism="Escherichia coli"
 /strain="1032 (Enterotoxigenic)"
 /db_xref="taxon:562"
 91.867
 /gene="Ltrh a subunit"
 91.867
 /gene="Ltrh a subunit"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin A subunit"
 /protein_id="BAA25725.1"
 /db_xref="GI:3062901"
 /translation="MKNITFFIFILLASPLYANGDKLYRADSRPDEIKRSGGLMPRG
 HNEYFDRTGOMINILYDHAQGTQGVRYDDGVSTSLRSAGHLAAGOSILSGSYTY
 IYVIATAPMFMVNDVLYGYSPHPYEQEYSAIGTYSQICMYRVNPGVIDERLHRN
 REYRDYRNILNIAEDGRIAGFPDPDQANREPMIHAAPGCGDSSRTITDTCN
 EETONLSTIYLKRIQSKVRQIFSDYQSEVDIYNRIDEL"
 864.1238
 /gene="Ltrh b subunit"
 864.1238
 /gene="Ltrh b subunit"
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin B subunit"
 /protein_id="BAA25726.1"
 /db_xref="GI:3062902"
 /translation="MNKVKCYVLTALSSLCAGAPOSITELCSFRMTQIYTIINDK
 ILTSYENAGKREMYITPKSGATFOVEPFGSHIDSOKAIEIRKMDITRLITLTETK
 IDKLGVNNKRTINSLAISMEN"
 BASE COUNT 488 a 258 c 286 g 402 t
 ORIGIN
 alignment_scores:
 Quality: 191.00 Length: 191
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 Alignment_block:
 US-09-528-682-1 x AB011677 ..
 Align seg 1/1 to: AB011677 from: 1 to: 1434
 5 LeuTYRtGAlaAsPserArGpROAsPGLuileYsArGserGlyGI 21
 |||||||
 157 TTATACGCTGCTGACTAGACCCCGAGATGAATAAACCTTCGCGAGG 206
 21 YLeuMetPROArGGLyHIsAnGLuTYrPheAsPArGGLyThrGlnMeTa 38
 |||||||
 207 TCTTATGCCAGAGCGCATATGAGTACTTCGATAGAGAACTCAATAGA 256
 38 snIleAnLeuTYrAsPHisAlArGGLyThrGlnThGlyPheValArG 54
 |||||||
 257 ATATTAATCTTATGATCAGCGAGAGAACACAAACGGCTTGTGCA 306
 55 TYrAsPAsPGLyTYrValSerThrSerLeuSerLeuArGserAlaHisLe 71
 |||||||
 307 TATGATGACGATATGTTTCACCTTCCTTACTTGTGAGAACTGCTCACTT 356
 71 AlaGLyGLInserIleLeuSerGlyTYrSerThrTYrTleTYrValI 88
 |||||||
 357 AGCAGGACAGCTATATTATCAGATATTCACCTTACTATATATATGTTA 406
 88 leaIaThrAlaProAsMerPheAsnValAsnAsPValleuGlyValTYr 104
 |||||||
 407 TAGGACAGCCCAATATGTTTAAATGTTAAATGATTAAGCGGATATAC 456
 105 SerProHsProTYrGLuGLInGluValSerAlaLeuGLyIleProTY 121
 |||||||

457 AGCCCTCACCATATGAACAGAGGTTTCGCTTAGGTGATACATA 506
 121 rSerGlnIleTYrGLyTYrTYrArGValAsnPhGlyValIleAspGLuA 138
 |||||||
 507 TTTCAGATATATGATGATGCTGCTTAAATTTGGCTTAATTATGATGAC 556
 138 rGLeuHisArGAsnArGGLuTYrArGAsPArGTYrTYrArGAsnLeuAsn 154
 |||||||
 557 GATTACATCGTACACGAGATATAGACCGGCTTTTACAGAAATCTGAT 606
 155 IleaIaProAlaGLuAsPGLyTYrArGleuAlaGLyPheProProAsP 171
 |||||||
 607 ATACTCGGCGAGAGATGTTACAGATATACAGTTTCCACCGGATCA 656
 171 sGLaIaTPArGGLuGLuProTPRIleHisAlaProGLInGLySG 188
 |||||||
 657 CCAAGCTTGAGAGAGAACCTGATTCATCAGACACACAGGTTGTG 706
 188 lYAsnSerSerArGThrIleThr 195
 707 GAAATTCATCAAGAACATTACA 729
 seq_name: gb_ba:ECOTOXA
 seq_documentation_block:
 LOCUS ECOTOXA 934 bp DNA linear BCT 26-APR-1993
 DEFINITION E.coli toxA gene encoding subunit A of heat-labile enterotoxin.
 ACCESSION K01995
 VERSION K01995.1 GI:148027
 KEYWORDS enterotoxin; heat-labile enterotoxin; toxA gene.
 SOURCE E.coli H10407 (serotype 078:H11) plasmid DNA, clone puy27.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 934)
 YAMAMOTO,T., TAMURA,T. and YOKOTA,T.
 Primary structure of heat-labile enterotoxin produced by
 Escherichia coli pathogenic for humans
 J. Biol. Chem. 259, 5037-5044 (1984)
 MEDLINE 84185610
 COMMENT This heat-labile enterotoxin is pathogenic for humans (Ltrh).
 Compared in [1] with LTP A and CT A (LTP-pathogenic for piglets,
 CT-cholera toxin).
 FEATURES
 source Location/Qualifiers
 1.934
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 158.934
 /note="heat-labile enterotoxin A prepeptide"
 /codon_start=1
 /transl_table=11
 /protein_id="AA24685.1"
 /db_xref="GI:148028"
 /translation="MKNITFFIFILLASPLYANGDKLYRADSRPDEIKRSGGLMPRG
 HNEYFDRTGOMINILYDHAQGTQGVRYDDGVSTSLRSAGHLAAGOSILSGSYTY
 IYVIATAPMFMVNDVLYGYSPHPYEQEYSAIGTYSQICMYRVNPGVIDERLHRN
 REYRDYRNILNIAEDGRIAGFPDPDQANREPMIHAAPGCGDSSRTITDTCN
 EETONLSTIYLKRIQSKVRQIFSDYQSEVDIYNRIDEL"
 158.211
 /note="heat-labile enterotoxin A signal peptide"
 212.787
 /product="heat-labile enterotoxin A1"
 794.931
 /product="heat-labile enterotoxin A2"
 165 c 186 g 285 t
 BASE COUNT 298 a 165 c 186 g 285 t
 ORIGIN 46 bp upstream of FokI recognition pattern.
 alignment_scores:
 Quality: 184.00 Length: 184
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOTOXA

Align seg 1/1 to: ECOTOXA from: 1 to: 934

```

5 LeuTYRARGALASPSEARGPPOPROASPGLUILEYSARGSERGLYGI 21
   |||||||
224 TTATACCGTGTGACTAGACCCCGAGATGAATAAAGCGTCCGAGG 273
   |||||||
21 YLeuMetProARGLylHisAnGIuTYRPhEaSPARGlyThrGlnMetA 38
   |||||||
274 TCTTATCCCGAGAGGCATATAGTACTTCGATAGAGAACTCAATGA 323
   |||||||
38 snlIeAsnLeuTYRAspHisAlaARGlyThrGlnThrGlyPheValArg 54
   |||||||
324 ATATTATCTTTATGATCAGCCGAGAGAACACAAACCGGCTTGTCAGA 373
   |||||||
55 TYTASpASpGlyTYRValSerThrSerLeuSerLeuArgSerAlaHisIe 71
   |||||||
374 TAGATGACGGATATGTTCCACTTCCTTAGTGTGAGAAAGTCTCACTT 423
   |||||||
71 uAlaGlyInSerIleLeuSerGIYTYRserThrTYRtyrIleTYRValI 88
   |||||||
424 AGCAGACAGCTATATTATCAGATATCCACTTACTATATATGTTA 473
   |||||||
88 leaIatHraIaProASmetPheAsnValAsnAspValLeuGlyValTYR 104
   |||||||
474 TAGGACAGACACCAATATGTTTAAATGTTAATGATGATTAGCGGTATAC 523
   |||||||
105 SerProHISProTYRGIuNGluValSerAlaLeuGIYleProTY 121
   |||||||
524 AGCCCTACCCCATATGAAACAGAGAGGTTTCGCGTAGGTGGAATACCAT 573
   |||||||
121 rSerGIuIleTYRGIYTYRtyrArgValAsnPhEgIYValIleAspGIuA 138
   |||||||
574 TTCTCAGATATATGATGATGATGCTGTTAATTTGGTGATGATGATGAC 623
   |||||||
138 rGLeuHISArgAsnArgGIuTYrArgAspArgTYTYRtyrArgAsnLeuAsn 154
   |||||||
624 GATTACATCGTAAACAGGAAATATAGACCCGCTATTACAGAAATCTGAAT 673
   |||||||
155 lIeAlaProAlaGIuAspGIYTYRArgLeuAlaGlyPheProProAspH 171
   |||||||
674 ATACCTCGGAGAGAGATGTTACAGATTAGCAGGTTTCCACCGGATCA 723
   |||||||
171 sGlnAlaTPArgGIuGIuProTrPIleHISHisAlaProGlnGlyCysG 188
   |||||||
724 CCAAGCTTGGAGAGAGAACCCGTGATTTCATCAGCACCAAGGTTGTG 773
   |||||||
188 lY 188
   ||
774 GA 775

```

seq_name: gb_ba:S60731

seq_documentation_block:

LOCUS S60731 1275 bp DNA linear BCT 29-JAN-2002

DEFINITION heat-labile enterotoxin A subunit, heat-labile enterotoxin B subunit [Escherichia coli, 21d, Genomic, 1275 nt].

ACCESSION S60731

VERSION S60731.1 GI:408994

KEYWORDS

SOURCE Escherichia coli 21d.

ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 1275)

AUTHORS Inoue,T., Tsuji,T., Koto,M., Iinamura,S. and Miyama,A.

TITLE Amino acid sequence of heat-labile enterotoxin from chicken enterotoxigenic Escherichia coli is identical to that of human strain H 10407

JOURNAL FEMS Microbiol. Lett. 108 (2), 157-161 (1993)

MEDLINE 93252225

PUBMED

8486242

REMARK

Genbank staff at the National Library of Medicine created this entry [NCBI glibseq 131491] from the original journal article. This sequence comes from Fig. 2.

FEATURES

Location/Qualifiers

source

1..1275

/organism="Escherichia coli"

/db_xref="taxon:562"

68..844

/note="This sequence comes from Fig. 2; conceptual translation presented here differs from translation in publication; LTC A subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin A subunit"

/protein_id="AAC60440.2"

/db_xref="GI:18390349"

/translation="MKNTFFIFILLASPIYANGDKLYRADRPDEIKRSGGLMPRG HNEFYDRGTOMKININDYHARGCTGTFYRDDGYSTSLSAHLAGSILSGSTYX IYVATAPNMFENVNDLVSPHPTFDEVSALGILYISQITGMTRVNFVYIDRLHRN REYRDYRYRNLIAPAEDEGYRLAGPPPHQAMREBPWLIHAPQGGGDSRRITGDTCN EETQWLSTIYLKTKYSKVRQIFSDYQSEVDIYNRIKREL"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig. 2; LTC B subunit"

/codon_start=1

/transl_table=11

/product="heat-labile enterotoxin B subunit"

/protein_id="AAC60441.1"

/db_xref="GI:408996"

/translation="MNKYCYLFTFALLSLICAGAPOSITELCEYRNTOIYTIINDK ILSTESMAGKREMYITPKSATQVEYPSQIHDSKKAIERKDLRIYLYLETK IDKLGVNKKTPNSTAISMEN"

841..1215

/note="This sequence comes from Fig.

121 rSerGlnIleTyrGlyTyrTyrArgValAsnPhgIleValIleAspGlu 138
 484 TTTCAGAAATATATAGTATGATGTTATTTGGTGTGATGATGAAAC 533
 138 rGluHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
 534 GATTACATCTTACAGAGGATATAGACCGGATTTACAGAAATCTGAAT 583
 155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspH 171
 584 ATAGCTCCGGCAGAGATGGTTACAGATTAGCAGGTTTCCACCGGATCA 633
 171 sGlnAlaTTPArgGluGluProTrrPheHisAlaProGlnGlyCysG 188
 634 CCAAGGTTGGAGAGAACCCGTGATTCATCATGCACCAAGGTTGTG 683
 188 Ly 188
 684 GA 685

seq_name: gb_pat:A04915

seq_documentation_block:

LOCUS A04915 777 bp DNA linear PAT 14-JUL-1993
 DEFINITION Artificial sequence LTA gene for labile toxin A.

ACCESSION A04915

VERSION A04915.1 GI:412519

KEYWORDS labile toxin A.

SOURCE synthetic construct.

ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 777)

AUTHORS Hayes, M.V., Harford, S. and Ross, G.W.

TITLE Microbiological process

JOURNAL Patent: EP 0145486-A 7 19-JUN-1985;

GLAXO GROUP LIMITED

FEATURES

source

1..777 Location/Qualifiers

/organism="synthetic construct"

/db_xref="taxon:32630"

1..777

/codon_start=1

/transl_table=1

/product="labile toxin A"

/protein_id="CA00403.1"

/db_xref="GI:412520"

/translation="MKNTTFIFTLASPLYANGDRLYRADSRPDEIKRSGGLMPRG

HNEYEDRGTOMNINLYDHARGTORGFEVYDDGYVTSLSRSHLAGOSITSGYSTYR

RYVATATNMFNVNDVAGVSPHREDEYSALGGIPYSQITGWIRVNFQVIDERLHRN

REYDRYRNINIAPEDEGYRLAGPPDHQAMREHPWIIHAPQCGGSSRTTGTGTCN

ETQNLSTIYREVSQSKYKRIQIFSDYOSEVDIYNIRDEL"

BASE COUNT 255 a 135 c 164 g 222 t

ORIGIN

alignment_scores:

Quality: 179.00 Length: 179

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A04915 ..

Align seg 1/1 to: A04915 from: 1 to: 777

62 ThSertLeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeu 78
 238 ACTTCTCTTATGTTTGAAGAGTCTCACTTACGACAGCTATATATATC 287
 78 rGlyTyrSerThrTyrTyrIleTyrValIleAlaThrAlaProAsnMetp 95
 288 AGGATATTCACCTTACATATATATGTTATAGGACAGCAACAAATATGT 337

95 heAsnValAsnAspValLeuGlyValTyrSerProHisProTyrGluGln 111
 338 TTATATGTTATGATGATATAGGGGTATACAGCCCTCACCATATATGAAC 387
 112 GluValSerIleLeuGlyGlyIleProTyrSerGlnIleTyrGlyTrrpy 128
 388 GAGGTTTCTCGTTAGGAGTAATACCATATCTCAGATATATAGGATGTA 437
 128 rArgValAsnPhgIleValIleAspGluArgLeuHisArgAsnArgGlu 145
 438 TCGTGTAAATTTTGGTGTGATGTATGACAGATTACATCATGACAGGAA 487
 145 rArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGly 161
 488 ATAGAGACCCGGTATACAGAAATCTGAAATATACCTCCGCGAGAGGAT 537
 162 TyrArgLeuAlaGlyPheProProAspHisGlnAlaTrrPargGluIupr 178
 538 TACAGATTACAGAGTTTCCACCGGATCACCAAGCTTGGAGAGAAAGACC 587
 178 OTrrPLeHisAlaProGlnGlyCysGlyAsnSerSerArgThrIlet 195
 588 CTGAGATTCATCATGCACACACAGGTTGTGAATTCATCAAGAAACATCA 637
 195 hrcGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIletTyr 211
 638 CAGGTGATCTTGTAAATGAGAGACCCAGAACTGTGACACAAATATATCTC 687
 212 ArgGluTyrGlnSerLysValLysArgGlnIlePheSerAspTrrGlnse 228
 688 AGGGAATATCAATCAAAAGTTTAAAGAGCAGACATTTTTCACATATCAGTC 737
 228 rGluValAspIleTyrAsnArgIleArgAspGluLeu 240
 738 AGAGGTGACATATATATACAGAAATTCGGATGAAATTA 774

seq_name: gb_ba:EMDLTA

seq_documentation_block:

LOCUS EMDLTA 723 bp DNA linear BCT 23-APR-1996

DEFINITION Plasmid EMD 299 (from E.coli) heat-labile enterotoxin subunit A

(LTA) gene, 3' end.

ACCESSION M57244.1 GI:531190

VERSION M57244.1 GI:531190

KEYWORDS enterotoxin; heat-labile enterotoxin.

SOURCE plasmid EMD 299 DNA.

ORGANISM plasmid EMD 299

REFERENCE 1 (bases 1 to 723)

AUTHORS Tsuji, T., Inoue, T., Miyama, A., Okamoto, K., Honda, T. and Mawatani, T.

TITLE A single amino acid substitution in the A subunit of Escherichia

coli enterotoxin results in a loss of its toxic activity

JOURNAL J. Biol. Chem. 265 (36), 22520-22525 (1990)

MEDLINE 91093102

COMMENT On Aug 20, 1994 this sequence version replaced gi:148531.

FEATURES

source

1..723 Location/Qualifiers

/organism="Plasmid EMD 299"

/plasmid="Plasmid EMD 299"

/specific_host="Escherichia coli"

/db_xref="taxon:2464"

1..723

/gene="LT-A"

<1..723

/gene="LT-A"

/codon_start=1

/transl_table=1

/product="heat-labile enterotoxin subunit A"

/protein_id="AAB59161.1"

/db_xref="GI:148532"

/translation="NGDRLYRADSRPPEIRKSGGLMPRGHNEYFDRTGTONINLYDH

AUTHORS
TITLE
IMMUNOGENIC DETOXIFIED MUTANTS OF CHOLERA TOXIN AND OF THE TOXIN
JOURNAL
PATENT: WO 9313202-A 42 08-JUL-1993;
FEATURES
source
Location/Qualifiers
1. .711
/organism="Escherichia coli"
/db_xref="taxon:562"
gene
1. .711
/gene="LT1 subunit A"
CDS
<1. .711
/gene="LT1 subunit A"
/codon_start=1
/transl_table=11
/protein_id="CAA01280.1"
/db_xref="GI:641009"
/translation="NGDRLYRADSRPDEIKRFRSLMRPNENYEDRGTOMINLYDHA
RGTOGEVRYDDGYVSTSLSLSAHLAQYLISGLTYIYIAMPVNVDIYSYP
HPYEEVSAIGIPYSOITGYWVRYNFGVIDERLHNRREYRNLINAPAEQYRL
AGFPDHOAMREEPWIIHAPGCGDSSRTITGDCNFTONLSTIYLREYOSKVKROI
FSDYQSEVDIYNRIDEL"
BASE COUNT 234 a 123 c 156 g 198 t
ORIGIN
alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x A16419 ..
Align seg 1/1 to: A16419 from: 1 to: 711
103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyIle 119
|||||
295 GTATACAGCCCTCACCATATGAACAGAGGTTTCGCTTAGTGGAAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrArgValAsnPhcGlyValIle 136
|||||
345 ACATATATTCACATATATGATGATGCTGTTAATTTGGTGATTC 394
|||||
136 spGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrArgAsn 152
|||||
395 ATGACGATTACATCTGTAACAGGAATATAGACCGGTAATACAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhePro 169
|||||
445 CTGAATATACCTCCGCGCAGAGATGTTACAGATTACAGGTTCCACAC 494
|||||
169 oAspHisGlnAlaTyrArgGluGluProTyrPheHisAlaProGln 186
|||||
495 GGATCACCACCAAGCTTGGAGAGAGAACCTGATTCATCATGCACACAG 544
|||||
186 lYcysGly 188
|||||
545 GTTGTGCA 552
seq_name: gb_pat:ARI18597
seq_documentation_block:
LOCUS ARI18597 711 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6149919.
ACCESSION ARI18597
VERSION ARI18597.1 GI:14100507
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 711)
AUTHORS Domenighini, M., Rappuoli, R., Pizza, M. and Hol, W.
TITLE Immunogenic detoxified mutants of cholera toxin and of the toxin

JOURNAL
PATENT: US 6149919-A 5 21-NOV-2000;
FEATURES
source
Location/Qualifiers
1. .711
/organism="unknown"
BASE COUNT 234 a 123 c 156 g 198 t
ORIGIN
alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x ARI18597 ..
Align seg 1/1 to: ARI18597 from: 1 to: 711
103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyIle 119
|||||
295 GTATACAGCCCTCACCATATGAACAGAGGTTTCGCTTAGTGGAAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrArgValAsnPhcGlyValIle 136
|||||
345 ACATATATTCACATATATGATGATGCTGTTAATTTGGTGATTC 394
|||||
136 spGluArgLeuHisArgAsnArgGluTyrArgAspArgTyrArgAsn 152
|||||
395 ATGACGATTACATCTGTAACAGGAATATAGACCGGTAATACAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPhePro 169
|||||
445 CTGAATATACCTCCGCGCAGAGATGTTACAGATTACAGGTTCCACAC 494
|||||
169 oAspHisGlnAlaTyrArgGluGluProTyrPheHisAlaProGln 186
|||||
495 GGATCACCACCAAGCTTGGAGAGAGAACCTGATTCATCATGCACACAG 544
|||||
186 lYcysGly 188
|||||
545 GTTGTGCA 552
seq_name: gb_ba:ECELTA
seq_documentation_block:
LOCUS ECELTA 868 bp DNA linear BCT 12-SEP-1993
DEFINITION E. coli gene eltA encoding the A subunit of the heat-labile
enterotoxin (LT).
ACCESSION V00275.1 GI:41339
VERSION V00275.1 GI:41339
KEYWORDS enterotoxin; signal peptide.
SOURCE
ORGANISM
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1 (bases 1 to 868)
AUTHORS Spicer, E.K. and Noble, J.A.
TITLE Escherichia coli heat-labile enterotoxin. Nucleotide sequence of
the A subunit gene
JOURNAL J. Biol. Chem. 257 (10), 5716-5721 (1982)
MEDLINE 82167425
FEATURES
source
Location/Qualifiers
1. .868
/organism="Escherichia coli"
/db_xref="taxon:562"
56. .>868
/product="messenger RNA"
90. .143
90. .854
sig_peptide
CDS
/codon_start=1
/transl_table=11
/product="enterotoxin subunit A"

/protein_id="CAA23532.1"
/db_xref="GI:41340"
/translation="MKNTFFIFILLASLYANGDRLYRADSRPPEDEIKRFSRLPRG
NEFDGTOMINILYDHDARGTQGFVRIDGIVTSLSLSAHLAQYLLSGYSLLTY
IYIANFENVNDIVSYSPHPYDEVSALGILYISQYMWRYNFGVIDERLHREYR
DHYRNLIAPAEDGYRLAGPPDHQAMREPEWIIHAPGCGDSSKTLTGDCNEETQ
NLSTYLRFGSKVRQIOPSDYQSEVDYINRIREDL"

BASE COUNT 284 a 142 c 179 g 263 t
ORIGIN

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x ECELTa ..

Align seg 1/1 to: ECELTa from: 1 to: 868

103 ValtyrSerProHisProtyrGluGlnGluValSerAlaLeuGlyIle 119
|||||
438 GTATACAGCCCTCACCATATGAAACAGAGAGGTTCTGCTAGGTTGGAAT 487
119 eProtyrSerGlnIleTyrglyTyrtprTyrtprValAspHeglyValIle 136
|||||
488 ACCATATTCAGATATATGATGATGATCGTTAAATTTGGTGTGATTG 537
136 spgIuAtyGleuHisArgAsnArgIuTyrtprArgAspArgtyrTyrtprAsn 152
|||||
538 ATGAACGATTCATCATCGTAACAGGAAATATAGACACCGTATTACAGAAAT 587
153 LeuAsnIleAlaProAlaGluAspGlyTyrtprGleuAlaGlyPheProPr 169
|||||
588 CTGAATATAGCTCCGCGACAGAGATGTTACAGATTACAGGTTCCACACC 637
169 OAspHisGlnAlaIleTprArgGluGluIupProTprIleHisAlaProGln 186
|||||
638 GCATCACCACCAACCTTGAGAGAGAACCTTGATTCATCATCACCACAG 687
186 IycysGly 188
|||||
688 GTTGTGGA 695

seq_name: gb_ba:ECU87

seq_documentation_block:
LOCUS ECU87 595 bp DNA linear BCT 28-OCT-1996
DEFINITION E.coli LT87 gene for heat-labile enterotoxin.
ACCESSION X83966
VERSION X83966.1 GI:1648865
KEYWORDS heat-labile enterotoxin; LT 87 gene.
SOURCE
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 595)
AUTHORS Germani,Y. and Desperrier,J.M.
TITLE Nucleotide sequence variations in genes encoding heat labile
enterotoxins of Escherichia coli isolated in South Pacific
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 595)
AUTHORS Germani,Y.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-1995) Y. Germani, Institut Pasteur, 28 rue du Dr
Roux, F-75727 Paris Cedex 15, FRANCE
COMMENT Related sequence: M15361-3 (Yakamoto).
FEATURES
Location/Qualifiers
1..595
/organism="Escherichia coli"
/strain="ETEC LT 87"

/db_xref="taxon:562"
gene 205..579
/gene="LT 87"
CDS 205..579
/gene="LT 87"
/codon_start=1
/transl_table=11
/product="heat-labile enterotoxin"
/protein_id="CAA58800.1"
/db_xref="GI:1648866"
/db_xref="SWISS-PROT:P13811"
/translation="MKVKFYVLFATLSSLCAGAPQSITELCEYHNTQIYIINDK
ILSYTSMAGKREMYITFKSGATFOVEVPSQIHDSOKAIEEMKDTLRITLYLETK
IDKLCVMNNKTPNSIAISMEN"

BASE COUNT 226 a 105 c 113 g 151 t
ORIGIN

alignment_scores:
Quality: 34.00 Length: 34
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x ECU87 ..

Align seg 1/1 to: ECU87 from: 1 to: 595

173 ALATprArgGluGluIupProTprIleHisAlaProGlnIlycysGlyAs 189
|||||
3 GCTTGGAGAGAGAACCCCTGATTCATCATCAGCACACAGGTTGTGGAA 52
189 nSerSerArgThrIleHcrglyAspThrCysAsnGluGluThrGlnAsn 206
|||||
53 TTCTATCAGACACATTCACGTGCTACTTGTATGAGAGAGACCAGATC 102
206 eu 206
||
103 TG 104

seq_name: gb_pat:A06082

seq_documentation_block:
LOCUS A06082 791 bp DNA linear PAT 25-MAY-1993
DEFINITION PERK620 HindIII-fragment.
ACCESSION A06082
VERSION A06082.1 GI:412718
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 791)
AUTHORS Ceska,M., Krizek,F., Hoesgenauer,G. and Besemer,J.
TITLE Plasmids and bacteria strains, and method of preparing the same
JOURNAL Patent: EP 0084522-A 1 27-JUL-1983;
SANDOZ AG; SANDOZ-PATENT-GMBH; SANDOZ-ERRINDUNGEN
Verwaltungsgesellschaft m.b.H
FEATURES
Location/Qualifiers
1..791
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 206 a 179 c 193 g 174 t 39 others
ORIGIN

alignment_scores:
Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A06082 ..

Align seg 1/1 to: A06082 from: 1 to: 791

198 ThrcysasnlgutThrglnasnleuSerThrIleTyrleuArgGluTyr 214
|||||
694 ACTTGTAATGAGGACCGACCAATCTGACGACATATATCTCAGGGAATA 743
214 rGlnSerLyValIysArgGlnIlePheSerAspTyrGln 227
|||||
744 TCAATCAAAAGTTAAGAGCGAGATATTTTCAGACTATCAG 783

seq_name: gb_pat:A06083

seq_documentation_block:

LOCUS A06083 791 bp DNA linear PAT 25-MAY-1993
DEFINITION PE8620 HindIII-fragment (Reverse complement).
ACCESSION A06083
VERSION A06083.1 GI:411208
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 791)
AUTHORS Ceska,M., Kricek,F., Hoegenauer,G. and Besemer,J.
TITLE Plasmids and bacteria strains, and method of preparing the same
JOURNAL Patent: EP 0084522-A 2 27-JUL-1983;
SANDOZ AG; SANDOZ-PATENT-GMBH; SANDOZ-ERFINDUNGEN
Verwaltungsgerichtsschaft m.b.H

FEATURES
Source Location/Qualifiers

1..791

/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 174 a 193 c 179 g 206 t 39 others
ORIGIN

alignment_scores:

Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A06083/rev ..

Align seg 1/1 to reverse of: A06083 from: 1 to: 791

198 ThrcysasnlgutThrglnasnleuSerThrIleTyrleuArgGluTyr 214
|||||
98 ACTTGTAATGAGGACCGACCAATCTGACGACATATATCTCAGGGAATA 49
214 rGlnSerLyValIysArgGlnIlePheSerAspTyrGln 227
|||||
48 TCAATCAAAAGTTAAGAGCGAGATATTTTCAGACTATCAG 9

seq_name: gb_pat:A02701

seq_documentation_block:

LOCUS A02701 795 bp DNA linear PAT 17-MAR-1993
DEFINITION Synthetic DNA sequence for enterotoxoid production.
ACCESSION A02701
VERSION A02701.1 GI:344645
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 795)
AUTHORS
TITLE NEW PLASMIDS AND BACTERIAL STRAINS AND PRODUCTION PROCESS THEREO F
JOURNAL Patent: WO 8302456-A 1 21-JUL-1983;
FEATURES Location/Qualifiers

Source

1..795

/organism="synthetic construct"
/db_xref="taxon:32630"

misc_feature 1..795

BASE COUNT 207 a /note="enterotoxoid and antigens production"
ORIGIN 180 c 193 g 174 t 41 others

alignment_scores:
Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A02701 ..

Align seg 1/1 to: A02701 from: 1 to: 795

198 ThrcysasnlgutThrglnasnleuSerThrIleTyrleuArgGluTyr 214
|||||
694 ACTTGTAATGAGGACCGACCAATCTGACGACATATATCTCAGGGAATA 743
214 rGlnSerLyValIysArgGlnIlePheSerAspTyrGln 227
|||||
744 TCAATCAAAAGTTAAGAGCGAGATATTTTCAGACTATCAG 783

seq_name: gb_pat:A02702

seq_documentation_block:

LOCUS A02702 795 bp DNA linear PAT 17-MAR-1993
DEFINITION Synthetic DNA sequence for enterotoxoid production (reverse complement).
ACCESSION A02702
VERSION A02702.1 GI:344646
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 795)
AUTHORS
TITLE NEW PLASMIDS AND BACTERIAL STRAINS AND PRODUCTION PROCESS THEREO F
JOURNAL Patent: WO 8302456-A 2 21-JUL-1983;
FEATURES Location/Qualifiers

Source

1..795

/organism="synthetic construct"
/db_xref="taxon:32630"

misc_feature 1..795

BASE COUNT 174 a /note="enterotoxoid and antigen production"
ORIGIN 193 c 180 g 207 t 41 others

alignment_scores:

Quality: 30.00 Length: 30
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A02702/rev ..

Align seg 1/1 to reverse of: A02702 from: 1 to: 795

198 ThrcysasnlgutThrglnasnleuSerThrIleTyrleuArgGluTyr 214
|||||
102 ACTTGTAATGAGGACCGACCAATCTGACGACATATATCTCAGGGAATA 53
214 rGlnSerLyValIysArgGlnIlePheSerAspTyrGln 227
|||||
52 TCAATCAAAAGTTAAGAGCGAGATATTTTCAGACTATCAG 13

seq_name: gb_ba:AF452584

seq_documentation_block:

LOCUS AF452584 777 bp DNA linear BCT 14-JAN-2002
DEFINITION Vibrio cholerae strain 1322-69 CtxA (ctxA) gene, complete cds.
ACCESSION AF452584

VERSION AF452584.1 GI:18141148
KEYWORDS
SOURCE Vibrio cholerae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.
TITLE Genetic Analysis of the Virulence Regions, CTX f prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Serogroup Strains of Vibrio cholerae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA
FEATURES
source 1.777
/organism="Vibrio cholerae"
/strain="1322-69"
/db_xref="taxon:666"
/note="identical sequence found in Vibrio cholerae strains
571-88 (serogroup O105) and 203-93 (serogroup O141)
serogroup: O37"
1.777
/gene="ctxA"
1.777
/gene="ctxA"
CDS 1.777
/gene="ctxA"
/codon_start=1
/transl_table=11
/product="CtxA"
/protein_id="AA160525.1"
/db_xref="GI:18141149"
/translation="MWKIIFFVFIFLSFSFSYANDDKLYRADSRPPEIKOSGLMPRG
ONEYFDRTGTONINLYDHAFCOTGFVRHDDGVSTISLSRAHLVGOTILSGHSTYY
IYVITAFPMFVNVDVLCAGSPHDEQVSAIGTIPYQIYCWYRVHFGVLDLOLHRN
RGYRDRTYSLNDIAPADGTYGAGFPERRAREREPWHHAPPGCGNAPRSSMSNTCD
EKTQSLGVKFLDEYOSKVKRQIFSSGYQSDIDYHNRIKDEL"
BASE COUNT 243 a 129 c 169 g 236 t
ORIGIN
alignment_scores:
Quality: 27.00 Length: 27
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AF452584 ..
Align seq 1/1 to: AF452584 from: 1 to: 777
28 AsnGluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHI 44
|||||
136 AATGAGTACTTTGACGAGGACTCAATGATATCAACCTTTATGATCA 185
44 sAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
186 TGCAGAGGAACCTCAGACGGGATTGTAGG 216
seq_name: gb_ba:AF463400
seq_documentation_block:
LOCUS AF463400 777 bp DNA linear BCT 31-JAN-2002
DEFINITION Vibrio cholerae strain 571-88 CtxA (ctxA) gene, complete cds.
ACCESSION AF463400
VERSION AF463400.1 GI:18448888
KEYWORDS
SOURCE Vibrio cholerae.

ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.
TITLE Genetic Analysis of the Virulence Regions, CTX f prophage and
Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic
Serogroup Strains of Vibrio cholerae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-2001) Department of Epidemiology and Preventive
Medicine, University of Maryland School of Medicine, and VA
Maryland Health Care System, The Warehouse at Camden Yards, 322 W.
Camden Street, Suite 675, Baltimore, MD 21201, USA
FEATURES
source 1.777
/organism="Vibrio cholerae"
/strain="571-88"
/db_xref="taxon:666"
/note="serogroup: O105"
1.777
/gene="ctxA"
1.777
/gene="ctxA"
CDS 1.777
/gene="ctxA"
/codon_start=1
/transl_table=11
/product="CtxA"
/protein_id="AA169944.1"
/db_xref="GI:18448889"
/translation="MWKIIFFVFIFLSFSFSYANDDKLYRADSRPPEIKOSGLMPRG
ONEYFDRTGTONINLYDHAFCOTGFVRHDDGVSTISLSRAHLVGOTILSGHSTYY
IYVITAFPMFVNVDVLCAGSPHDEQVSAIGTIPYQIYCWYRVHFGVLDLOLHRN
RGYRDRTYSLNDIAPADGTYGAGFPERRAREREPWHHAPPGCGNAPRSSMSNTCD
EKTQSLGVKFLDEYOSKVKRQIFSSGYQSDIDYHNRIKDEL"
BASE COUNT 243 a 129 c 169 g 236 t
ORIGIN
alignment_scores:
Quality: 27.00 Length: 27
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AF463400 ..
Align seq 1/1 to: AF463400 from: 1 to: 777
28 AsnGluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHI 44
|||||
136 AATGAGTACTTTGACCGAGGACTCAATGATATCAACCTTTATGATCA 185
44 sAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
186 TGCAGAGGAACCTCAGACGGGATTGTAGG 216
seq_name: gb_ba:AF463401
seq_documentation_block:
LOCUS AF463401 777 bp DNA linear BCT 31-JAN-2002
DEFINITION Vibrio cholerae strain 203-93 CtxA (ctxA) gene, complete cds.
ACCESSION AF463401
VERSION AF463401.1 GI:18448890
KEYWORDS
SOURCE Vibrio cholerae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 777)
AUTHORS Li,M., Chen,Y., Kotetishvili,M., Stine,O.C., Morris,J.G. Jr.,
Sulakvelidze,A. and Sozhamannan,S.

TITLE Genetic Analysis of the Virulence Regions, CTX ϕ prophage and Vibrio Pathogenicity Island (VPI), in Diverse, Non-epidemic Serogroup Strains of Vibrio cholerae

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 777)

AUTHORS Li, M., Chen, Y., Kotelishvili, M., Stine, O. C., Morris, J. G. Jr., Sulakvelidze, A. and Sozhamannan, S.

TITLE Direct Submission

JOURNAL Submitted (27-DEC-2001) Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, and VA Maryland Health Care System, The Warehouse at Camden Yards, 322 W. Camden Street, Suite 675, Baltimore, MD 21201, USA

FEATURES

source 1..777

/organism="Vibrio cholerae"

/strain="203-93"

/db_xref="taxon:666"

/note="serogroup: O141"

1..777

/gene="ctxA"

1..777

/gene="ctxA"

/codon_start=1

/transl_table=1

/product="CtxA"

/protein_id="AAL69945.1"

/db_xref="GI:18448891"

/translation="MKILVFVFIFLSSFSYANDDKLYRADSRPPDEIKOSGGLMREG QNEYFDRTQMNIINLYDHARGTGTGVRHDDGVSTISLSRSHLVGOTILSGHSTYY IYVIATAPNMFVNVDVLAGYSPHDEQVYSAIGIPYSOIGYRVHFGVLDLQHRN RGYRDRYVNLDIAPADGGLAGFPPEHRAMREPEWIIHAPPGCGNAPRSSMSNTCD EKTQSLGVYFLEDEYQSKVKRQIFSGVQSDIDTHNRKDEL"

BASE COUNT 243 a 129 c 169 g 236 t

ORIGIN

alignment_scores:

Quality: 27.00 Length: 27

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AF463401 ..

Align seg 1/1 to: AF463401 from: 1 to: 777

28 AsnGluTyrPheaspArgGlyThrGlnMetasnIleasnLeuTyrAspHi 44

|||||

136 AATGAGTACTTTGACCGAGGACTCAATGAAATCAACCTTTATGATCA 185

44 sAlaArgGlyThrGlnThrGlyPheValArg 54

|||||

186 TGCAGAGGAGACTCAGACGGGATTGTAGG 216

seq_name: gb_ba:VIBCTXABA

seq_documentation_block:

LOCUS VIBCTXABA 1372 bp DNA linear BCT 01-FEB-2000

DEFINITION Vibrio cholerae ctxa gene and ctxb gene for cholera toxins, complete cds.

ACCESSION D30052

VERSION D30052.1 GI:487332

KEYWORDS ADP-ribosyltransferase; cholera toxin.

SOURCE Vibrio cholerae (sub-species O37, strain S7) (library: plasmid)

DNA, clone pK7310.

ORGANISM Vibrio cholerae

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE 1 (bases 1 to 1372)

AUTHORS Yamamoto, K., Do, V. G., Xu, M., Iida, T., Mawatani, T., Albert, M. J. and Honda, T.

TITLE Comparison of cholera toxin genes (ctxAB) of non-O1 vibrio strains 854 (O139-bengal) and S7 (O37) from two outbreaks unpublished (1994)

JOURNAL

REFERENCE 2 (bases 1 to 1372)

AUTHORS Yamamoto, K.

TITLE Direct Submission

JOURNAL Submitted (29-APR-1994) Kolchiro Yamamoto, Osaka University, Institute for Microbial Diseases, Department of Bacterial Infections; 3-1 Yamadaoka, Suita, Osaka 565, Japan (Tel:06-879-4066)

COMMENT Submitted (29-Apr-1994) to DDBJ by: Kolchiro Yamamoto

Institute for Microbial Diseases

Osaka University

3-1 Yamadaoka

Suita, Osaka 565

Japan

phone: 06-879-4066

Fax: 06-879-8277.

FEATURES

source 1..1372

/organism="Vibrio cholerae"

/strain="S7"

/sub_species="O37"

/db_xref="taxon:666"

/clone_lib="plasmid"

113..889

/gene="ctxA"

113..889

/gene="ctxA"

/codon_start=1

/transl_table=1

/product="cholera toxin"

/protein_id="BA06288.1"

/db_xref="GI:808898"

/translation="MKILVFVFIFLSSFSYANDDKLYRADSRPPDEIKOSGGLMREG QNEYFDRTQMNIINLYDHARGTGTGVRHDDGVSTISLSRSHLVGOTILSGHSTYY IYVIATAPNMFVNVDVLAGYSPHDEQVYSAIGIPYSOIGYRVHFGVLDLQHRN RGYRDRYVNLDIAPADGGLAGFPPEHRAMREPEWIIHAPPGCGNAPRSSMSNTCD EKTQSLGVYFLEDEYQSKVKRQIFSGVQSDIDTHNRKDEL"

BASE COUNT 449 a 217 c 269 g 437 t

ORIGIN

sig_peptide

/gene="ctxA"

886..1260

/gene="ctxB"

886..1260

/gene="ctxB"

/codon_start=1

/transl_table=1

/product="cholera toxin"

/protein_id="BA06289.1"

/db_xref="GI:808900"

/translation="MIKLKGVFVTLSSAVAHGTPTDLCAGYHNTQIHTLNDK ILSYFSLAGNREMAITTFKNGATTFQVEVPGSCHIDSQKAIERMKDTLRIVLFEAK VEKLCVNRNKKTPHAIATAISMAN"

886..948

/gene="ctxB"

217 c 269 g 437 t

alignment_scores:

Quality: 27.00 Length: 27

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIBCTXABA ..

Align seg 1/1 to: VIBCTXABA from: 1 to: 1372

28 AsnGluTyrPheaspArgGlyThrGlnMetasnIleasnLeuTyrAspHi 44

|||||

248 AATGAGTACTTTGACCGAGGACTCAATGAAATCAACCTTTATGATCA 297

44 sAlaArgGlyThrGlnThrGlyPheValArg 54

|||||

298 TGCAGAGAACTCAGACGGGATTGTTAGG 328

seq_name: gb_pat:A16422

seq documentation block:

LOCUS A16422 723 bp DNA linear PAT 06-JUL-1995

DEFINITION A subunit of cholera toxin CT.

ACCESSION A16422

VERSION A16422.1 GI:641010

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

IMMUNOGENIC DETOXIFIED MUTANTS OF CHOLERA TOXIN AND OF THE TOXIN

LT, THEIR PREPARATION AND THEIR USE FOR THE PREPARATION OF VACCINES

Patent: WO 9313202-A 45 08-JUL-1993;

Location/Qualifiers

1..723

/organism="Vibrio cholerae"

/db_xref="taxon:666"

1..723

/gene="CT subunit A"

<1..723

/codon_start=1

/protein_id="CAA01281.1"

/translation="NDKLIYRADSRPPEIKOSGLMPGSEYFDKGTOMNINLYD

ARGTGTGFVRHDDGYVSTISLSAHVGGTLLSGHSTYIVIAATAPMFWNDVLG

AVSPHDEVSALGIPYSQIYGVYVHFGVLDLHNRGVRDRYVSNLIDIAAAD

GYGLGFPPEHRAAREEPIHHAAPGCGNAPRSSISNCDERTQSIGVFLDEYQSKV

KQIFSGYOSDIDTNNRIKDEL"

BASE COUNT 227 a 125 c 163 g 208 t

ORIGIN

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A16422 ..

Align seg 1/1 to: A16422 from: 1 to: 723

seq_name: gb_pat:AR118598

seq documentation block:

LOCUS AR118598 723 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 7 from patent US 6149919.

ACCESSION AR118598

VERSION AR118598.1 GI:14100508

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

Immunogenic detoxified mutants of cholera toxin and of the toxin

LT, their preparation and their use for the preparation of vaccines

Patent: US 6149919-A 7 21-NOV-2000;

JOURNAL

FEATURES

source

BASE COUNT 227 a 125 c 163 g 208 t

ORIGIN

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AR118598 ..

Align seg 1/1 to: AR118598 from: 1 to: 723

seq_name: gb_ba:VIBCTXA

seq documentation block:

LOCUS VIBCTXA 777 bp DNA linear BCT 26-APR-1993

DEFINITION Vibrio cholerae enterotoxin A1 peptide (ctxA) gene, 5' end.

ACCESSION K02679

VERSION K02679.1 GI:155159

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

sig_peptide

CDS

mat_peptide

BASE COUNT 237 a 119 c 164 g 257 t

ORIGIN

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x VIBCTXA ..

Align seg 1/1 to: VIBCTXA from: 1 to: 777

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
280 GAGTACTTTCACGAGCTACTCAATGATATACACCTTATGATCATGCG 329
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
330 AAGAGGAACTCAGACGCGATTGTTCAG 357

seq_name: gb_pat:E00132

seq_documentation_block: 1148 bp DNA linear PAT 29-SEP-1997
LOCUS E00132
DEFINITION DNA coding of cholera toxin.
ACCESSION E00132
VERSION E00132.1 GI:2168431
KEYWORDS JP 1983222033-A/1.
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1148)
AUTHORS Arufuocru, N. and Do, B.M.
TITLE SUBUNIT A AND B OF DNA ARRANGEMENT, RECOMBINED DNA, CHOLERA TOXIN
AND MEDICINE
PATENT: JP 1983222033-A 1 23-DEC-1983;

JOURNAL

COMMENT
SMITH KLINE RIT
OS cholera(Vibrio)

PN JP 1983222033-A/1
PD 23-DEC-1983
PF 23-MAY-1983 JP 1983091416
PR 24-MAY-1982 US 82 381083
PI ARUFUOCRU NIJIERU, DO BURUDO MISHIERU
PC A61K39/106,A61K39/00,C07H21/04,C12N15/00;
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=ATCC 39050;
FH Key Location/Qualifiers

FT CDS 1..777
FT /product='subunit A of cholera toxin' FT
FT mat_peptide 1..774
FT /product='subunit A of cholera toxin' FT CDS
FT 774..1148
FT /product='subunit B of cholera toxin' FT
FT mat_peptide 774..1145
FT /product='subunit B of cholera toxin' FT

FEATURES
source Location/Qualifiers
1..1148
/organism='unidentified'
/db_xref='taxon:32644'

BASE COUNT 382 a 184 c 235 g 347 t
ORIGIN

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x E00132 ..

Align seg 1/1 to: E00132 from: 1 to: 1148

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTCACGAGCTACTCAATGATATACACCTTATGATCATGCG 188

45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGAGGAACTCAGACGCGATTGTTCAG 216

seq_name: gb_ba:VCTOXAB21

seq_documentation_block:

LOCUS VCTOXAB21 1362 bp DNA linear BCF 23-SEP-1991
DEFINITION V.cholerae genes for toxin proteins A and B, strain 2125.
ACCESSION X58786
VERSION X58786.1 GI:48420
KEYWORDS cholera toxin.
SOURCE Vibrio cholerae.
ORGANISM Vibrio cholerae.

REFERENCE 1 (bases 1 to 1362)
AUTHORS Dams,E.
TITLE Direct Submission

JOURNAL

Submitted (25-MAR-1991) E. Dams, Laboratory of Human Biochemistry,
RUGA University Antwerp, Gruenendurgerlaan 171, B-2020 Antwerp,
Belgium

REFERENCE 2 (bases 1 to 1362)
AUTHORS Dams,E., De Wolf,M. and Dierick,W.
TITLE Correction of the cholera toxin nucleotide sequence of the Vibrio
cholerae eltor strain 2125
JOURNAL Unpublished

FEATURES
source Location/Qualifiers

1..1362
/organism='Vibrio cholerae'
/strain='eltor 2125'
/db_xref='taxon:666'
/clone='PRT10824'
72..92
/rpt_type='TANDEM'
110..121
136..142
169..945
/codon_start=1
/transl_table=11
/product='cholera toxin A protein (CTA)'
/protein_id='CAA41592.1'
/db_xref='GI:48421'
/db_xref='SWISS-PROT:P01555'

/translation='MKITFVFPIFLSSFSYANDKLYRADSRPPPEIKQSGLMRG
QSEYEDRGTMNINLYDHARGTQTVRHHDDGVSTLSKSHLHGTTLSGHSTY
IYVATAPNMFVNDVLAGVSPHPDQVSLAGIGPYSGIYGRVHFGVLEQHLRN
RGVDRRYENLDIAPADGCGIAGFPPEHRAWNEEPWIHHAPGCGNADRSSMSNTCD
EKQSGVKEFLDEYQSKVRQIFSGVQSDIDTNRKIDEL'
169..222
223..942
/product='cholera toxin A protein (CTA)'
942..1004
942..1316
/codon_start=1
/transl_table=11
/product='cholera toxin B protein (CTB)'
/protein_id='CAA41593.1'
/db_xref='GI:48422'
/db_xref='SWISS-PROT:P01556'
/translation='MIKKEGVFETVLLSSAVAHGTPONTDLCAEYHNTQITLNDK
IFSYPESLAGKREMAITPFGKAIFQVEYPSGCHIDSOKKAIERMDTIRIVLTPEAK
VERLCVWNNKTRPHALAIASMAN'

BASE COUNT 455 a 205 c 276 g 426 t
ORIGIN

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VCTOXB21 ..

Align seg 1/1 to: VCTOXB21 from: 1 to: 1362

29 GluTyrPheaspArgGlyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
 ||||||||||||||||||||||||||||||||||||||||||||||||
 307 GAGTACTTTCACCGAGTACTCAATGATATCAACCTTATGATCATGC 356
 45 aArgGlyThrGlnThrGlyPheValArg 54
 ||||||||||||||||||||||||||||||||||||||||||||||||
 357 AAGAGAACTCAGACGGGATTGTAGG 384

seq_name: gb_ba:VIBCTYXAB

seq_documentation_block:

LOCUS VIBCTYXAB 1369 bp DNA linear BCT 01-FEB-2000
 DEFINITION Vibrio cholerae ctxA gene and ctxB gene for cholera toxins,
 complete cds.
 ACCESSION D30053
 VERSION D30053.1 GI:487333
 KEYWORDS ADP-ribosyltransferase; cholera toxin.
 SOURCE Vibrio cholerae (subspecies O139-Bengal, strain 1854) (library:
 plasmid) DNA, clone pKY340.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE 1 (bases 1 to 1369)
 Yamamoto, K., Do, Y.G., Xu, M., Iida, T., Miwatani, T., Albert, M.J. and
 Honda, T.
 TITLE Comparison of cholera toxin genes (ctxAB) of non-O1 vibrio cholerae
 strains 854 (O139-Bengal) and 57 (O37) from two outbreaks
 JOURNAL Unpublished (1994)
 REFERENCE 2 (bases 1 to 1369)
 Yamamoto, K.

JOURNAL
AUTHORS

COMMENT

Direct Submission
 Submitted (29-APR-1994) Koichiro Yamamoto, Osaka University,
 Institute for Microbial Diseases, Department of Bacterial
 Infections; 3-1 Yamadaoka, Suita, Osaka 565, Japan
 (Tel:06-879-4066)
 Submitted (29-APR-1994) to DDBJ by:
 Koichiro Yamamoto
 Institute for Microbial Diseases
 Osaka University
 3-1 Yamadaoka
 Suita, Osaka 565
 Japan
 Phone: 06-879-4066
 Fax: 06-879-8277.
 FEATURES
 source
 1..1369
 Location/Qualifiers
 /organism="Vibrio cholerae"
 /strain="1854"
 /sub_species="O139-Bengal"
 /db_xref="taxon:666"
 /clone_lib="plasmid"
 110..886
 /gene="ctxA"
 110..886
 /gene="ctxA"
 /codon_start=1
 /transl_table=11
 /product="cholera toxin"
 /protein_id="BAA06290.1"
 /db_xref="GI:808901"
 /translation="MKITIFVFIFLSEFSYANDKLYRADSRPPDEIKQSGGLMPRG
 OSEYFDRGTQNNINLYDHARGTQTFVRRHDDGYSTSLRSALHVGQTILSGHSTYY
 IYVITAPNMFVNVDLYGASPHDPEQESALGIPYISQYGMVHVHFGVLEQLHRN
 RGYRDYRYSNLDIAPAAGYGLAGFPPEHRAREPPIHAPPGCGNAPRRSSMSTCD
 EKTOSLGVKFLDEVGSKVKRQIFSGYQSDIDITHNRKDEL"
 110..163
 sig_peptide
 /gene="ctxA"
 883..1257

CDS

/gene="ctxB"
 883..1257
 /gene="ctxB"
 /codon_start=1
 /transl_table=11
 /product="cholera toxin"
 /protein_id="BAA06291.1"
 /db_xref="GI:808902"
 /translation="MKIKFGVFVLLSAYAHGTQNTDLCAEYHNTOIYTLNDK
 IFSYREIAGKREMAITFEKCAIFQVPEPSQIHDSQKAIERNKDTLRILAYLLEAK
 VEKLCVNNKTPHAIATISMAN"
 883..945
 /gene="ctxB"

sig_peptide

BASE COUNT 449 a 214 c 269 g 437 t
 ORIGIN

alignment_scores:

Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIBCTYXAB ..

Align seg 1/1 to: VIBCTYXAB from: 1 to: 1369

29 GluTyrPheaspArgGlyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
 ||||||||||||||||||||||||||||||||||||||||||||||||
 248 GAGTACTTTCACCGAGTACTCAATGATATCAACCTTATGATCATGC 297
 45 aArgGlyThrGlnThrGlyPheValArg 54
 ||||||||||||||||||||||||||||||||||||||||||||||||
 298 AAGAGAACTCAGACGGGATTGTAGG 325

seq_name: gb_ba:VCTOXB56

seq_documentation_block:

LOCUS VCTOXB56 1397 bp DNA linear BCT 30-JUN-1993
 DEFINITION V. cholerae genes for toxin proteins A and B, strain 569B.
 ACCESSION X58785 S55782
 VERSION X58785.1 GI:48888
 KEYWORDS cholera toxin.
 SOURCE Vibrio cholerae.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE 1 (bases 1 to 1397)

Dams, E.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAR-1991) E. Dams, Laboratory of Human Biochemistry,
 RUCA University Antwerp, Gruenenburgerlaan 171, B-2020 Antwerp,
 Belgium

REFERENCE 2 (bases 1 to 1397)

Dams, E., De Wolt, M. and Dierick, W.
 TITLE Nucleotide sequence analysis of the CT operon of the Vibrio
 cholerae classical strain 569B
 JOURNAL Biochim. Biophys. Acta 1090 (1), 139-141 (1991)
 MEDLINE 91355224
 FEATURES
 source
 1..1397
 Location/Qualifiers
 /organism="Vibrio cholerae"
 /strain="classical strain 569B"
 /db_xref="taxon:666"
 /clone="PJM17"
 72..127
 /rpt_type=TANDEM
 repeat_region
 -35_signal
 -10_signal
 CDS
 145..177
 171..177
 204..980
 /codon_start=1
 /transl_table=11
 /product="cholera toxin A protein (CTA)"
 /protein_id="CAA41590.1"


```
/db_xref="GI:4889"
/db_xref="SWISS-PROT:P01555"
/translation="WVKIIFVFIFLSFSYANDKLYRADSRPPEIKQSGCLMPRG
QSEYFDRTGMNINLYDHARGTGTGVRHDDGVSTSLRSANHLVGOTILSGHSTY
IYVATAPNMFNVNDVLGAYSPHDEQVSAAGIPIYQIYGMRYVHFGLDEQLHRN
RGYRDRYXSNLDIAPADGAYGACGPPERRHARREPMWIIHAPGCGNAPRRSSMTCD
EKTQSLGVKFLDEYQSKVKROIFSGYQSDIDITHNKDEL"
204..257
sig_peptide
258..977
mat_peptide
sig_peptide
977..1039
sig_peptide
977..1351
/codon_start=1
/translation="cholera toxin A protein (CTA)"
/product="cholera toxin A protein (CTA)"
/transl_table=11
/product="cholera toxin B protein (CTB)"
/protein_id="CAA41591.1"
/db_xref="GI:48890"
/db_xref="SPTREMBL:Q57193"
/translation="MIKLKFGVFTVLLSSAYAHGPNQITDLCAEYHNTQIHTLNDK
IESYTESLAGKREMAIITPKNGATFOVEPVSQIHDSQKKAERKKDILRIAYLLEAK
VERLCVWNKKTPHAIATISMAN"
1040..1348
mat_peptide
1040..207 c 281 g 449 t
BASE COUNT 460 a
ORIGIN

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x VCTOXAB56 ..
Align seg 1/1 to: VCTOXAB56 from: 1 to: 1397

29 GUTYrPhespargglyThGImetCasnIIeasneutyraSPHISAL 45
|||||
342 GAGTACTTACCGAGGACTCTCAATGAAATATCAACCTTATGATCATGC 391
|||||
45 AARGlyThrGlnThrGlyPheValArg 54
|||||
392 AAGAGGAACCTCAGACGGGATTGTGTTAGG 419
|||||

seq_name: gb_ba:VCCTX

seq_documentation_block:
LOCUS VCCTX 2020 bp DNA linear BCT 30-MAR-1995
DEFINITION Vibrio cholera toxin (ctx) operon DNA sequence from strain 2125.
ACCESSION X00171
VERSION X00171.1 GI:48347
KEYWORDS overlapping genes: signal peptide; tandem repeat.
SOURCE Vibrio cholerae.
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 2020)
AUTHORS Mekalanos,J.U., Swartz,D.J., Pearson,G.D., Harford,N., Groyne,F.
and de Wilde,M.
TITLE Cholera toxin genes: nucleotide sequence, deletion analysis and
vaccine development
JOURNAL Nature 306 (5943), 551-557 (1983)
MEDLINE 84068199
FEATURES
Location/Qualifiers
1..2020
/organism="Vibrio cholerae"
/db_xref="taxon:666"
misc_feature 419..425
/note="tandem repetitive sequence"
misc_feature 426..432
/note="tandem repetitive sequence"
misc_feature 433..439
/note="tandem repetitive sequence"
```

```
promoter 457..468
/note="35 region"
promoter 483..489
/note="10 region"
RBS 504..508
/note="Shine-Dalgarno sequence"
RBS 507..510
/note="Shine-Dalgarno sequence"
CDS 516..1292
/note="ctx A"
/codon_start=1
/transl_table=11
/protein_id="CAA24995.1"
/db_xref="GI:48348"
/db_xref="SWISS-PROT:P01555"
/translation="WVKIIFVFIFLSFSYANDKLYRADSRPPEIKQSGCLMPRG
QSEYFDRTGMNINLYDHARGTGTGVRHDDGVSTSLRSANHLVGOTILSGHSTY
IYVATAPNMFNVNDVLGAYSPHDEQVSAAGIPIYQIYGMRYVHFGLDEQLHRN
RGYRDRYXSNLDIAPADGAYGACGPPERRHARREPMWIIHAPGCGNAPRRSSMTCD
EKTQSLGVKFLDEYQSKVKROIFSGYQSDIDITHNKDEL"
516..524
sig_peptide
522..1151
mat_peptide
1152..1289
mat_peptide
1289..1663
gene
1289..1663
/gene="ctx B"
1289..1351
sig_peptide
1289..1351
/gene="ctx B"
1289..1663
CDS
1289..1663
/gene="ctx B"
/codon_start=1
/transl_table=11
/protein_id="CAA24996.1"
/db_xref="GI:758351"
/db_xref="SWISS-PROT:P01556"
/translation="MIKLKFGVFTVLLSSAYAHGPNQITDLCASHNTQIHTLNDK
IESYTESLAGKREMAIITPKNGATFOVEPVSQIHDSQKKAERKKDILRIAYLLEAK
VERLCVWNKKTPHAIATISMAN"
1352..1660
mat_peptide
1352..1660
/gene="ctx B"
/product="ctx B"

BASE COUNT 602 a 352 c 446 g 620 t
ORIGIN

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x VCCTX ..
Align seg 1/1 to: VCCTX from: 1 to: 2020

29 GUTYrPhespargglyThGImetCasnIIeasneutyraSPHISAL 45
|||||
654 GAGTACTTGTGACCGAGGACTCTCAATGAAATATCAACCTTATGATCATGC 703
|||||
45 AARGlyThrGlnThrGlyPheValArg 54
|||||
704 AAGAGGAACCTCAGACGGGATTGTGTTAGG 731
|||||

seq_name: gb_pat:E03512

seq_documentation_block:
LOCUS E03512 2020 bp DNA linear PAT 29-SEP-1997
DEFINITION Cholera toxin gene.
ACCESSION E03512
VERSION E03512.1 GI:2171728
KEYWORDS JP 199209488-A/17.
SOURCE synthetic construct.
```


ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 2020)
AUTHORS Shirai,H.
TITLE PRIMER FOR AMPLIFICATION OF GENE
JOURNAL Patent: JP 1992099488-A 17 31-MAR-1992;
UNITIKA LTD
COMMENT PN JP 1992099488-A/17
PD 31-MAR-1992
PP 20-AUG-1990 JP 1990219523
PI SHIRAI HIROMASA
PC C12N15/11,C12Q1/68;
CC strandedness: Double;
CC topology: Linear;
FH Key location/Qualifiers
FT misc-feature 1..2020
location/Qualifiers
1..2020
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 602 a 352 c 446 g 620 t
ORIGIN
alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x E03512 ..
Align seg 1/1 to: E03512 from: 1 to: 2020
29 GtutyrpheasparglythrGlnmetasnleasnlenturyrasphisal 45
|||||
654 GAGTACTTGGACCGAGTACTCAATGAAATGATCACTTTATGATCATGC 703
45 aArgglythrGlnthrGlyphevalArg 54
|||||
704 AAGAGGAACTCAGACGGAGTTGTTAGG 731
seq_name: gb_ba:AF175708
seq_documentation_block:
LOCUS AF175708 3480 bp DNA linear BCT 15-NOV-1999
DEFINITION Vibrio cholerae accessory cholera enterotoxin (ace), zonular
occludens toxin (zot), cholera enterotoxin A-subunit (ctxA), and
cholera enterotoxin B-subunit (ctxB) genes, complete cds.
ACCESSION AF175708
VERSION AF175708.1 GI:5764645
KEYWORDS
SOURCE Vibrio cholerae.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE 1 (bases 1 to 3480)
AUTHORS Shin,H.J., Park,Y.C. and Kim,Y.C.
TITLE Cloning and nucleotide sequence analysis of the virulence gene
JOURNAL cassette from Vibrio cholerae KNIH002 isolated in Korea
REFERENCE 2 (bases 1 to 3480)
AUTHORS Misaiumuhag Hoij1 35 (3), 205-210 (1999)
TITLE Shin,H.J., Park,Y.C. and Kim,Y.C.
JOURNAL Direct Submission
Submitted (04-AUG-1999) School of Life Sciences, Chungbuk National
University, Kaesin-Dong san 48, Cheongju, Chungbuk-do 361-763,
South Korea
FEATURES
source 1..3480
location/Qualifiers
1..3480
/organism="Vibrio cholerae"
/strain="KNIH002"
/db_xref="taxon:666"

gene /country="South Korea"
287..577
/gene="ace"
CDS 287..577
/gene="ace"
/codon_start=1
/transl_table=11
/product="accessory cholera enterotoxin"
/protein_id="AAD51357.1"
/db_xref="GI:5764645"
/translation="MAMDLYDLIDFTNLYIKGLIMIESKIFVIOFWEWSQKV
IDMFITYPLIQDAIDMLPPQYSGFLFGLDQALIVLQALMTLRPALNALM."
574..1773
/gene="zot"
574..1773
/gene="zot"
/note="cholera enterotoxin"
/codon_start=1
/transl_table=11
/product="zonular occludens toxin"
/protein_id="AAD51358.1"
/db_xref="GI:5764647"
/translation="MSIFIHGAPSGSYKTSQALMRLPAIKSGRHITNRYGINTLR
MAKYLKMDVSDISIEFIDTDPDGRUTNAREWMAKDAFLFIDCGIWPRLTVN
LKALDTPDVLVADRESEFVAFDMHRHGWIDICLTENIAKVNMMIREAEIGRHF
NRATVGLGAKFTLTTHDAANSQOMDSHALTRQVKRIPSTIFKMYASTTGAKARDYAG
TALMKDRKILFLEGVWFLMEYSFYGLNDNPFTFGNDATIESEQSEQSAATVNAV
GSKAVAPASFGFCIGRICVODGEFVVGDERRLVNDIPYRGIMATGHIIYKDTLV
PRTSGSVPTPLFASRYRYKVLPLPDENHPVPTFAQALMVEYKRGKLPKTEWDK
KCLNLSIF"
1872..2648
/gene="ctxA"
1872..2648
/gene="ctxA"
/codon_start=1
/transl_table=11
/product="cholera enterotoxin A-subunit"
/protein_id="AAD51359.1"
/db_xref="GI:5764648"
/translation="MWKIIFFVFIFLSSFSFYANDDKLYRADSPPEIDKOSGLMPRG
QSEVFRGROMNINLYDARCGTQGFVRRHDDGYSTISIRSAHLVGGTILSHSTY
ITYIATAPMENVNDVLCATSPHPDEQESALGCIPTYSQIGMYRVHVGVDQAHNRN
KGYRDYYSNLDIAPADGGLAGFPPEHRAWREPPWJHHAAPGCGNAPRSMSTCD
EKTQSLGKVFIDEYQSKYKROIIFSGYSDIDTHRIKDEL"
2645..3019
/gene="ctxB"
2645..3019
/gene="ctxB"
/codon_start=1
/transl_table=11
/product="cholera enterotoxin B-subunit"
/protein_id="AAD51360.1"
/db_xref="GI:5764649"
/translation="MIKIKGVFPYVILSSAAGHTPONIDICAEVHNTQIHITLNDK
ITFSTESLAGKREMAIITFKNGATFOVEPGSQIHDSOKAIDRKADTLRLIAYLTLEAK
VERLCVWNKRTPHALIAAISMAN"
BASE COUNT 948 a 721 c 808 g 1003 t
ORIGIN
alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AF175708 ..
Align seg 1/1 to: AF175708 from: 1 to: 3480
29 GtutyrpheasparglythrGlnmetasnleasnlenturyrasphisal 45
|||||
2010 GAGTACTTGGACCGAGAACCTCAATGAAATGATCACTTTATGATCATGC 2059

45 aargglythrGlnThrGlyPheValArg 54
 |||||
 2060 AAGAGGAACTCAGACGCGATTGTTAGG 2087

seq_name: gb_ph:AF414369

seq_documentation_block:

LOCUS AF414369 4275 bp DNA linear PHG 02-OCT-2001
 DEFINITION Vibrio phage CTX Ace (ace), Zot (zot), CtxA (ctxA), and CtxB (ctxB)
 genes, complete cds.
 AF414369
 ACCESSION AF414369.1 GI:15825396

KEYWORDS

SOURCE
 ORGANSIM

Vibrio phage CTX.
 Viruses; ssDNA viruses; Inoviridae; Inovirus.

REFERENCE
 1 (bases 1 to 4275)
 Bhattacharyya,T., Nandy,R.K. and Nair,G.B.
 The entire core region of the ctx-phi (ctx-prophage) in VCE 232, an
 environmental strain of V. cholerae

REFERENCE
 2 (bases 1 to 4275)
 Bhattacharyya,T., Nandy,R.K. and Nair,G.B.

AUTHORS
 Direct Submission

JOURNAL
 Submitted (27-AUG-2001) Microbiology, NCCED, Calcutta, India

FEATURES

Location/Qualifiers

1..4275
 /organism="Vibrio phage CTX"

/specific_host="Vibrio cholerae strain VCE 232"

/db_xref="taxon:141904"

/note="prophage: core region"

1543..1833

/gene="ace"

1543..1833

/gene="ace"

/codon_start=1

/transl_table=1

/product="ace"

/protein_id="AAL09683.1"

/db_xref="GI:15825396"

/translation="MLMMDLYDMLIDGFTWLVIKLGIMIESKIFVIOFWMSQKV
 IDMFYIPLIOAIDMLPPQYSGFLFGLDQALAIYQALMTREPALALNL"

1830..3029

/gene="zot"

1830..3029

/gene="zot"

/codon_start=1

/transl_table=1

/product="zot"

/protein_id="AAL09684.1"

/db_xref="GI:15825397"

/translation="WKRIIFYFIFLSSFYANDKLYLRASRPPEDEIKSGCGLMPG
 QSEYFDKOTONINLYDHARGTGTGFVHHGQVSTISLSAHLVGGTILSGHSTYY
 IYVIAIAPNMENVDVLAAYSPHPDEQVSAIGIIPYQIIGVYRVHGVILDEQLHRN
 RGYRDYYSNIDIAADGVLGAGFPREHRAWRBEPWIHAPPCGNGAPRSSMSNTCD

gene

CDS

EKTOSLGVFLDEYOSKVKROIFSGYQSDIDTNRKIDEL"
 3901..4275
 /gene="ctxB"
 3901..4275
 /gene="ctxB"
 /codon_start=1
 /transl_table=1
 /product="ctxB"
 /protein_id="AAL09682.1"

BASE COUNT 1132 a 904 c 997 g 1242 t

ORIGIN

alignment_scores:

Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AF414369 ..

Align seg 1/1 to: AF414369 from: 1 to: 4275

29 GltYrPheaspArgglyThrGlnMetAsnIleAsnLeuTyraPheValArg 45

3266 GAGTACTTGTACCGAGCTACTCAATGAATCAACCTTATGATCATGTC 3315

45 aargglythrGlnThrGlyPheValArg 54

3316 AAGAGGAACTCAGACGCGATTGTTAGG 3343

seq_name: gb_ba:AE004224

seq_documentation_block:

LOCUS AE004224 10646 bp DNA linear BCT 31-JUL-2000

DEFINITION Vibrio cholerae chromosome I, section 132 of 251 of the complete

chromosome.

ACCESSION AE004224 AE003852

VERSION AE004224.1 GI:9655952

KEYWORDS

SOURCE

ORGANSIM

Vibrio cholerae.

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE

AUTHORS

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

Medical Center Dr. Rockville, MD 20850, USA

1..10646

Location/Qualifiers

/organism="Vibrio cholerae"

/strain="N16961"

gene
/serotype="O1"
/db_xref="taxon:666"
/chromosome="1"
/note="biotype: E1_Tor"
complement(215. .439)
/gene="VC1452"
complement(215. .439)
/gene="VC1452"
/note="similar to GP:2564354; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="RstC protein"
/protein_id="AAF94609.1"
/db_xref="GI:9655953"
/translation="MSKLVNTRMDVYDSIEDLNNAALYLRSGAYTDEIAHQYNLICD
KIIDOGIYNTFRSLPSLNPKALTEPSL"
complement(333. .907)
/gene="VC1453"
complement(533. .907)
/gene="VC1453"
/note="similar to GP:2564353; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="RstB1 protein"
/protein_id="AAF94610.1"
/db_xref="GI:9655954"
/translation="MMKLMVNMKSRFVYFGASHSEGVSKTAPLIPIVLYPKRIRQ
MKNDGQCITREGLQHOEVKVFSSDAMTRKLEOTAPVLTFFNEPDEPDSNLIYD
QVVCSLFDVNPVGKRLDKPKNT"
complement(885. .1964)
/gene="VC1454"
complement(885. .1964)
/gene="VC1454"
/note="similar to GP:2564352; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="RstA1 protein"
/protein_id="AAF94611.1"
/db_xref="GI:9655955"
/translation="MKQLEFIDEIQLDTNAPVFDYLANSPYASFRHAKSDLS
SLIMAPLPKPYRMAKRTPEQKEKLELYKOKNVAMMERLEVCLEHVLGRSPMRDK
GLGYGNSCHLSMYSKHKHGVVALGNNNTCYFOIEGVGCRVLEHTSILFELHMLD
LIGCYSLRIDLAVDDPFGLEFGREYAKKAYSDAPRTAAGRAPNGGEVLEPEPKI
INSEFEVGSRESRIYRYNKAAGLGLDMHMRNREVELKMDP IDVLNTEGVFAGICA
YASIIINSIPKVVYTKRROYALDIHSRIKMAROVKITLFDISKHFGDLERVFAGLI
SKEIHDDSLNLPDSTMKLIDETMGD"
2090. .2428
/gene="VC1455"
2090. .2428
/gene="VC1455"
/note="similar to GP:2564351; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="transcriptional repressor RstR"
/protein_id="AAF94612.1"
/db_xref="GI:9655956"
/translation="MKIKERLANOKRAINKTQAOAMADEIGISLYSKYSEGGLPMT
ENLVKADALEISIDELCGRMATDENOEMLRLKIQOQDEDEOKAISMVLESLMRH
SRSKILNHGA"
complement(3030. .3404)
/gene="VC1456"
complement(3030. .3404)
/gene="VC1456"
/note="similar to GB:K01170 SP:P01556 GB:D30052 GB:D30053
GB:X58785; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cholera enterotoxin, B subunit"
/protein_id="AAF94613.1"

gene
/db_xref="GI:9655957"
/translation="MIKIKGCVFFTVLLSSAYAHGTPOINTDLCAEYHNTQIYTLNDK
IFSYTESLAGKREMAITTEKNGALFOVEVPGSIDSQKAIEMKDLRLIAYLTLEAK
VEKICVWNKTRPHAIAMISMAN"
complement(3401. .4177)
/gene="VC1457"
complement(3401. .4177)
/gene="VC1457"
/note="similar to SP:P01555 GB:D30052 GB:D30053 GB:K02679
GB:X58785; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cholera enterotoxin, A subunit"
/protein_id="AAF94614.1"
/db_xref="GI:9655958"
/translation="MWKIIFFVFTEFLSESYANDDKLYRADSRPEDEIKOSGLMPRG
OSEYFDRGCTQKNINLYDARGTQGFVNRHDDGYVTSISRLAHLVGOTILSGSTY
IYVIRAPNMFNRNDVIGANISPHPEDEQVSNLGGITPISQIYGMRYRFGVLDPELHRN
RGYRDRYSLNDIAPADGYGLAGPPEHRARREPMTHARPCCGNAPRSSMSNTCD
EKTQSLGVKFLDEYOSVKRQIFSGYOSDIDITHRIRDEL"
complement(4276. .5475)
/gene="VC1458"
complement(4276. .5475)
/gene="VC1458"
complement(4276. .5475)
/note="similar to GB:M83563 SP:P38442 PID:155315;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="zona occludens toxin"
/protein_id="AAF94615.1"
/db_xref="GI:9655959"
/translation="MSIFIHGARGSYKTSGLMWLRLLPAIKSGRHITTYNGINLER
MAKYLMKDVSDISIEFIDTDHPDRLTWARPMHARKDAFLFIDECGRIMPRRLTVN
LKALDTPDLVAEDRESEFEVAFDMHRHMGDICTLPINAKVNMIRAEALISYRHF
NRATYVGAGKFTITTHDANGSDSHALTQOVKIGSPITKMYASTTGKARDTMG
TALKDKRIILFELGVPLMFSESYTGCHLNDIFPGGDATIESQSEQSKATYGMNV
GSKAVAPASFGFCIGRLCVODGFVTVGERIRLVDNDIDIPGIMAGHNIYKOTLY
FEETESGVPTELFASSYRYKVLPLDPNHRFVDTFAQAQLWVEVRKGLPIKTENDK
KGLNSIF"
complement(5472. .5765)
/gene="VC1459"
complement(5472. .5765)
/gene="VC1459"
/note="similar to SP:P38441 GB:Z22569 PID:298062;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="accessory cholera enterotoxin"
/protein_id="AAF94616.1"
/db_xref="GI:9655960"
/translation="MMLMDITLYMMLIDGFTWLYIKGIMNIESKIFVIOFWMESQK
VIDMFTIYPLIQOALIDMLPOYSGFLFFGLDQALAIYQALMTFALRALNL"
complement(5762. .7045)
/gene="VC1460"
complement(5762. .7045)
/gene="VC1460"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF94617.1"
/db_xref="GI:9655961"
/translation="MAGSILMSLGHNYRLCVGLIFGLREGESESPFMRYFLPLITLLF
LSPSYTASAINCDPMTTSHQDLFGFSGPIYOSVLPDCMDIDKDDGYFVMSCLSN
NGDYCKGLXKRPESGVGPPNMPMDLISGASAEICITVPCPGECECVLPPEPDSPY
DGLSSRSKSAFNQYTKRNOSEWASTLNHVSQVSHSODMVOANTFPHADRYLESTAYN
NRLGQDMYELEIRIDVDTQREVRKADELYSRVAASVTVLSELVNLVAIDELKD
SLGTVYVPPNDOPNPTPPDSSSPYTGALNTISKLNLTLETISQOQDTNMTALSGRC
SNPCPCPIREAEETELETAAQONKONKINERTISQALHQFGSAVVSFCSYVEAFGY
NLCFDFSLSENIHITIMVILAMAVIILAAILLFR"
complement(7056. .7304)
/gene="VC1461"
complement(7056. .7304)
/gene="VC1461"


```

/gene="VC1461"
/Note="similar to ; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="colonization factor"
/protein_id="AAF94618.1"
/db_xref="GI:9655962"
/translation="MESSKKKLNFKSTLSGVFLPSAFANQALAAADGLVTEVT
KTLGTSKDTVIALGPIMGVGAIVLIVTGLIRKAK"
complement(7440..7823)
/gene="VC1462"
complement(7440..7823)
/Note="similar to GP:2564358; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="RatB2 protein"

```

alignment_scores:
 Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AE004224/rev ..

Align seg 1/1 to reverse of: AE004224 from: 1 to: 10646

```

29 GUTYrPheasparGlyThrGlnMetAsnIleAsnLeuTyraSPHISAI 45
|||||
4039 GAGTACTTTCACGAGGTACTCAATGATATCACTTTATGATCATGC 3990
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
3989 AAGAGCACTCAGACGGATTGTAGG 3962

```

seq_name: gb_ba:ECOELTA2

seq_documentation_block:

LOCUS ECOELTA2 165 bp DNA linear BCT 26-APR-1993

DEFINITION E.coli heat labile enterotoxin subunit a gene; a1/a2 junction.

ACCESSION K00433

VERSION K00433.1 GI:145829

KEYWORDS enterotoxin.

SOURCE Escherichia coli DNA.

ORGANISM Escherichia coli

REFERENCE 1 (bases 1 to 165)
 Spitzer,E.K., Kavanaugh,W.M., Dallas,W.S., Falkow,S.,
 Konigsberg,W.H. and Schaefer,D.E.
 Sequence homologues between a subunits of Escherichia coli and
 Vibrio cholerae enterotoxins
 Proc. Natl. Acad. Sci. U.S.A. 78, 50-54 (1981)

JOURNAL MEDLINE 81233767

COMMENT see other loci beginning <ecoelt>.

FEATURES

source

1..165

/organism="Escherichia coli"

/db_xref="taxon:562"

BASE COUNT 59 a 30 c 35 g 41 t

ORIGIN about 300 bp after seg 1.

alignment_scores:
 Quality: 24.00 Length: 24
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOELTA2 ..

Align seg 1/1 to: ECOELTA2 from: 1 to: 165

```

195 ThrGlyAspThrCysAsnGluThrGlnAsnLeuSerThrIleTyrlle 211
|||||
55 ACAGGTATACCTTGTATGAGAGAGACCAGAAATCTGAGACAAATATATCT 104
211 uArgGlyThrGlnSerIleVal 218
|||||
105 CAGGGAATATCAATCAAAAGTT 126

```

seq_name: gb_ba:ECOELT

seq_documentation_block:

LOCUS ECOELT 774 bp DNA linear BCT 25-JUN-1996

DEFINITION E.coli heat labile enterotoxin a 3' end and enterotoxin b (tox b)

ACCESSION J01646

VERSION J01646.1 GI:145828

KEYWORDS enterotoxin.

SOURCE Escherichia coli (strain H10407) DNA.

ORGANISM Escherichia coli

REFERENCE 1 (bases 190 to 292)

AUTHORS Yamamoto,T., Tamura,T., Ryoji,M., Kaji,A., Yokota,T. and Takano,T.

TITLE Sequence analysis of the heat-labile enterotoxin subunit B gene

J. Bacteriol. 152 (1), 506-509 (1982)

J. Bacteriol. 152 (1), 506-509 (1982)

83007048

2 (bases 1 to 228)

Yamamoto,T., Tamura,T., Yokota,T. and Takano,T.

Overlapping genes in the heat-labile enterotoxin operon originating

from Escherichia coli human strain

Mol. Gen. Genet. 188 (2), 356-359 (1982)

83114628

3 (bases 190 to 774)

Yamamoto,T. and Yokota,T.

Sequence of heat-labile enterotoxin of Escherichia coli pathogenic

for humans

J. Bacteriol. 155 (2), 728-733 (1983)

83265593

see also other loci beginning <ecoelt>.

COMMENT

FEATURES

source

1..774

/organism="Escherichia coli"

/strain="H10407"

/specific_host="Homo sapiens"

/db_xref="taxon:562"

<1..208

/codon_start=2

/transl_table=11

/product="enterotoxin A"

/protein_id="AA02981.1"

/db_xref="GI:1395121"

/translation="AMREPMTHHAPQGGSSRTITGPTCNBERQNLSTIYLKRYQS

KYKRIETSPDQSEVDIYNRIKREL"

195..197

/gene="tox b"

195..579

/gene="tox b"

205..267

/gene="tox b"

205..579

/gene="tox b"

/codon_start=1

/transl_table=11

/product="enterotoxin B"

/protein_id="AA02982.1"

/db_xref="GI:1395122"

/translation="MNKRCYVLFALLSSICAYGAPQSITEICSEYRNPTQITINDK

ILSYTESMAGKREMYITPFKSGATFOVEVPGSOHIDSOAKAIEBMKDTLRLITYLLETGK

IDKICVMNKNKTPNSIAAISMEN"

CDs

sig_peptide

195..579

195..579

205..267

205..579

mat_peptide 268 .576
 /gene="toxh"
 /product="enterotoxin B"
 BASE COUNT 286 a 142 c 153 g 193 t
 ORIGIN

alignment_scores:
 Quality: 24.00 Length: 24
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x ECOELT ..

Align seg 1/1 to: ECOELT from: 1 to: 774

214 TyrglnSerLysValLysArgGlnIlePheSerAspTyrGlnSerGluVa 230
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 125 TATCAATCAAAAGTTAAAGAGCAGATATTTCAGACTATCAGTCAGAGCT 174
 230 LAspIleTyrAsnArgIleArg 237
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 175 TGACATATATACAGAAATCGG 196

seq_name: gb_ba:ECOPHLEGS

seq_documentation_block:

LOCUS ECOPHLEGS 900 bp DNA linear BCT 26-APR-1993
 DEFINITION Escherichia coli heat-labile enterotoxin gene, 5' cds.
 ACCESSION M61015
 VERSION M61015.1 GI:147190
 KEYWORDS heat-labile enterotoxin.
 SOURCE E.coli, DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1 to 900)
 Trachman,J.D. and Maas,W.K.
 TITLE Characterization of the control region of the porcine heat-labile
 enterotoxin in Escherichia coli
 JOURNAL Unpublished (1991)
 FEATURES
 source
 location/Qualifiers
 1..900

CDs
 mRNA
 /organism="Escherichia coli"
 /strain="K-12"
 /db_xref="taxon:562"
 722..>900
 /product="heat-labile enterotoxin"
 781..>900
 /codon_start=1
 /transl_table=11
 /product="heat-labile enterotoxin"
 /protein_id="AA24335.1"
 /db_xref="GI:147191"
 /translation="MKNIFFIFILLASPLVANGDRLYRADSRPDEIRSGSL"

BASE COUNT 227 a 213 c 181 g 279 t
 ORIGIN

alignment_scores:
 Quality: 22.00 Length: 22
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOPHLEGS ..

Align seg 1/1 to: ECOPHLEGS from: 1 to: 900

1 AsngLysArgLeuTyrArgAlaSerArgProProAspGluIleIly 17
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 835 AATGGCAGACATATATACCGTGTACTGTAGACCCCGATGAATAATA 884

17 sargSerGlyGlyLeu 22
 ||||||||||||||||
 885 ACGTTCGGAGAGTCTT 900

seq_name: gb_ba:VIBCTA1

seq_documentation_block:

LOCUS VIBCTA1 255 bp DNA linear BCT 26-APR-1993
 DEFINITION Vibrio cholerae cholera gene: subunit A alpha-chain 5' end.
 ACCESSION J01846
 VERSION J01846.1 GI:155156
 KEYWORDS cholera toxin.
 SEGMENT 1 of 2
 SOURCE Vibrio cholerae (strain El Tor 1621) DNA.
 ORGANISM Vibrio cholerae.

REFERENCE 1 (bases 1 to 255)
 Gennaro,M.L. and Greenaway,P.J.
 TITLE Nucleotide sequences within the cholera toxin operon
 JOURNAL Nucleic Acids Res. 11, 3855-3861 (1983)
 MEDLINE 83246519
 COMMENT clean copy provided by authors.
 FEATURES
 source
 location/Qualifiers
 1..255

BASE COUNT 85 a 23 c 38 g 109 t
 ORIGIN

alignment_scores:
 Quality: 13.00 Length: 13
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIBCTA1 ..

Align seg 1/1 to: VIBCTA1 from: 1 to: 255

5 LeuTyrArgAlaSerArgProProAspGluIleLys 17
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 199 TTATATCGCGCAGATCTCTAGACCTCTGTAGTAATAAATAAG 237

seq_name: gb_ba:VIBCTA2

seq_documentation_block:

LOCUS VIBCTA2 251 bp DNA linear BCT 26-APR-1993
 DEFINITION Vibrio cholerae cholera toxin gene: subunit A gamma-chain 3' end &
 subunit B 5' end.
 ACCESSION J01847
 VERSION J01847.1 GI:155157
 KEYWORDS cholera toxin.
 SEGMENT 2 of 2
 SOURCE Vibrio cholerae (strain El Tor 1621) DNA.
 ORGANISM Vibrio cholerae.

REFERENCE 1 (bases 1 to 251)
 Gennaro,M.L. and Greenaway,P.J.
 TITLE Nucleotide sequences within the cholera toxin operon
 JOURNAL Nucleic Acids Res. 11, 3855-3861 (1983)
 MEDLINE 83246519

FEATURES
 source
 location/Qualifiers
 1..251

BASE COUNT 96 a 38 c 38 g 79 t
 ORIGIN

alignment_scores:
 Quality: 12.00 Length: 12

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIRBTA2 ..

Align seg 1/1 to: VIRBTA2 from: 1 to: 251

213 GluTyrGlnSerLysValLysArgGlnIlePheSer 224
|||||
61 GAATACCAATCTAAAGTTAAAGACAATATTTTCA 96

seq_name: gb_ba:VIBTOX

seq_documentation_block:

LOCUS VIRBTOX 512 bp DNA linear BCT 26-APR-1993
DEFINITION Vibrio cholerae toxA and toxB genes for cholera enterotoxin

ACCESSION K01170.1 GI:155296

VERSION enterotoxin.

KEYWORDS Vibrio cholerae (El Tor biotype strain 62746) DNA, clone pCVD002.

SOURCE Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

ORGANISM 1 (bases 1 to 512)

REFERENCE Lockman,H. and Kaper,J.B.

AUTHORS Nucleotide sequence analysis of the A2 and B subunits of Vibrio

TITLE cholerae enterotoxin

JOURNAL J. Biol. Chem. 258, 13722-13726 (1983)

MEDLINE 84061784

COMMENT The termination codon of subunit A2 overlaps the initiation codon

of subunit B by 4 bp. The correct translation of the 2 subunits is

the result of a shift in reading frames at this junction. Previous

protein structural analysis has shown that the A2 subunit is bound

to a precursor of the A subunit. Therefore, the 'met' at position 1

is not an initiator.

FEATURES Location/Qualifiers

source 1..512

/organism="Vibrio cholerae"

/db_xref="taxon:666"

1..141

/note="enterotoxin subunit A2"

/codon_start=1

/transl_table=11

/protein_id="AA027572.1"

/db_xref="GI:155296"

/translation="MIKLFQVFVTLSSAYAHGTPONTDLCAEYHNTQIYTLNDK

IESYTESLAGKREMAITFEKSGTEVEVPGSOHDSOKKAIERMKDRLRIYLVLEAK

VEKLCVMNKKTPHAIATISMAN"

141..200

/note="enterotoxin subunit B signal peptide"

201..509

/product="enterotoxin subunit B mature peptide"

BASE COUNT 199 a 74 c 88 g 151 t

ORIGIN 188 bp upstream of NdeI site.

alignment_scores:

Quality: 12.00 Length: 12

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x VIRBTOX ..

Align seg 1/1 to: VIRBTOX from: 1 to: 512

213 GluTyrGlnSerLysValLysArgGlnIlePheSer 224
|||||
55 GAATACCAATCTAAAGTTAAAGACAATATTTTCA 90

seq_name: gb_ba:ECOELTBP

seq_documentation_block:

LOCUS ECOELTBP 604 bp DNA linear BCT 16-APR-1996

DEFINITION E.coli (from pig) heat-labile enterotoxin subunit B gene (eltB),

complete cds.

ACCESSION M17873.1 GI:145832

VERSION M17873.1

KEYWORDS eltB gene; enterotoxin.

SOURCE Escherichia coli DNA.

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Escherichia.

REFERENCE 1 (bases 1 to 407)

AUTHORS Dallas,W.S. and Falkow,S.

TITLE Amino acid sequence homology between cholera toxin and Escherichia

coli heat-labile toxin

JOURNAL Nature 288 (5790), 499-501 (1980)

MEDLINE 81074965

AUTHORS 2 (bases 17 to 604)

REFERENCE Leong,J., Vinal,A.C. and Dallas,W.S.

TITLE Nucleotide sequence comparison between heat-labile toxin B-subunit

cistrons from Escherichia coli of human and porcine origin

INFECTION. Immun. 48 (1), 73-77 (1985)

MEDLINE 85156481

COMMENT [2] reverses [1].

A long inverted repeat which could be a termination signal is

located at positions 533-564.

FEATURES Location/Qualifiers

source 1..604

/organism="Escherichia coli"

/specific_host="Sus scrofa"

/db_xref="taxon:562"

33..407

/gene="eltB"

33..95

/gene="eltB"

33..407

/gene="eltB"

/note="precursor"

/codon_start=1

/transl_table=11

/product="enterotoxin"

/protein_id="AA098065.1"

/db_xref="GI:145833"

/translation="MNKVKCYLFTALSSLYAHGAPOTTTELCSERYRNTQIYTLNDK

ILSYTESMAGKREMYITFEKSGTEVEVPGSOHDSOKKAIERMKDRLRIYLVLETK

IDKLCVMNKKTPHAIATISMAN"

96..404

/gene="eltB"

/product="enterotoxin"

114

/gene="eltB"

/note="Nature 288, 499-501 (1980)"

/citation=[1]

223

/gene="eltB"

/note="Nature 288, 499-501 (1980)"

/citation=[1]

362

/gene="eltB"

/note="Nature 288, 499-501 (1980)"

/citation=[1]

BASE COUNT 225 a 112 c 116 g 151 t

ORIGIN 17 bp upstream of EcoRI site.

alignment_scores:
Quality: 11.00 Length: 11
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x ECOELTBP ..

Align seg 1/1 to: ECOELTBP from: 1 to: 604

230 ValaspietTyrasnArgilearAspGluLeu 240

1 GTTGACATATATACAGATTGGGATGAATTA 33

seq_name: gb_ba:ECOETOXHL

seq_documentation_block:

LOCUS ECOETOXHL 1262 bp DNA linear BCT 18-NOV-1994

DEFINITION Escherichia coli (strain 41) heat-labile enterotoxin type IIB

ACCESSION M28523.1 GI:576584

KEYWORDS LT-IIB gene; enterotoxin type IIB.

SOURCE Escherichia coli (strain 41) DNA.

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Escherichia.

REFERENCE 1 (bases 1 to 1262)
Pickert,C.L., Twiddy,E.M., Coker,C. and Holmes,R.K.

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

gene

-35_signal

-10_signal

CDS

sig_peptide

mat_peptide

sig_peptide

CDS

seq_documentation_block:

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

/protein_id="AAA53286.1"
/db_xref="GI:576586"
/translation="MSFKIIRKAFVIMATLVSOAHGASQFQDNCRRTASLVEGV
ELTKYISDINNNDGMYVSSGCVWRISRAKDDPDVMTAEMRKIAAAVILSGMRVN
MCASPSSPNYIMALELAE"

mat_peptide
963.1259
/gene="LT-IIB"
/note="B chain of heat-labile enterotoxin type IIB"
/product="enterotoxin"

BASE COUNT 369 a 195 c 285 g 413 t

ORIGIN

alignment_scores:

Quality: 11.00 Length: 11
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x ECOETOXHL ..

Align seg 1/1 to: ECOETOXHL from: 1 to: 1262

38 AsnIleasnLeuTyrAspHisAlaArgGlyThr 48

278 AACATTAACTATATGACATGACGAGTACT 310

seq_name: gb_pat:AR019034

seq_documentation_block:

LOCUS AR019034 42 bp DNA linear PAT 05-DEC-1998

DEFINITION Sequence 14 from patent US 5783196.

ACCESSION AR019034

VERSION AR019034.1 GI:3974148

KEYWORDS

SOURCE

ORGANISM

Unknown.

REFERENCE 1 (bases 1 to 42)

AUTHORS

TITLE

JOURNAL

Patent: US 5783196-A 14 21-JUL-1998;

Location/Qualifiers

1.42

/organism="unknown"

BASE COUNT 14 a 8 c 10 g 10 t

ORIGIN

alignment_scores:

Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AR019034 ..

Align seg 1/1 to: AR019034 from: 1 to: 42

192 ArgThrIleThrGlyAspThrCysAsnGlu 201

13 AGAACAAATTAACAGGTGATCTGTAAATGAG 42

seq_name: gb_pat:A69697

seq_documentation_block:

LOCUS A69697 2724 bp DNA linear PAT 07-MAY-1999

DEFINITION Sequence 15 from Patent WO9807864.

ACCESSION A69697

VERSION A69697.1 GI:4774310

KEYWORDS

SOURCE

unidentified.

ORGANISM

unidentified.


```

unclassified.
1 (bases 1 to 2724)
REFERENCE 1 Foster K.A., Quinn, C.P. and Shone, C.C.
AUTHORS RECOMBINANT TOXIN FRAGMENTS
TITLE Patent: WO 9807864-A 15 26- FEB-1998;
JOURNAL FOSTER KEITH ALAN (GB)
FEATURES
source Location/Qualifiers
1..2724
/organism="unidentified"
/db_xref="taxon:32644"
1..2721
/product="unnamed"
1..2724
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAB42522.1"
/db_xref="GI:4774311"
/translation="MQFVNKQFNYKDPVNGVDIAYIKIIPNAGOMPVYAKIHNKIMV
IPERDTFTNPEEGDLNPPPEAKQVPSYSDSYLSTDNKQNYLGVYKLFPRITSD
LGRMLTSTVIRGIPWGGSTIDTELKVIDTNCINVIOPGSGVRSSEINLVITGPSADI
IOFECKSRGHEVLTNRNGYSTOYIRSPDTPRFEESELEVDITPBLGACFATDPA
VTLAHELHAGHRLGIAINPVRKVTNMAIYEMSGLEVESEELRTEGHDAKTIDS
LOENEFRLYYNKKFKDIASITLNAKSIYGTASIQMKVFEKYLLESDTGKFSVD
KLKFDKLYKMLTEIYEDNPFVFEKVLNKRITLNFDAVKFINIYKVTYIYDGFNL
RTNLANPNGONTFETNNMNFETLKNFTGLCFEYKILCYRGITTSKTSLDGYNKIE
GRCDGALNDLCIKVNMWDLFESSEDNFTNDLNGKEELTSDNINIAEENISLIDLIO
VYTPNPENPENISIEUNSSDIQOLELMPRIEKPNGKKELDKTYTFHILRAQEF
EHGKSRILITNSVNEALNPSRVYTFESSDYKAKYKATKTEAMFLGWELVYDFDE
TSEVSTDKIADITIIIPYIGALNIGMLYDDVGALIFSGAVILFEIPEIAIPV
LTSFALVSYIANKVLTVQIDIALSKRNKMDVYKVIYTNMLAVNTQIDILIRKMK
EALENOAMATKAIINYQYNOYEEENNNINIDLSKLNSEKNNKAMINIKPLNOC
SVSYLNMASNIPYGVKREDEPDLAKLYIDNRGTLIGOVDRKDKVNTLSTDI
PFQISKYVDNRLLSTFTFYIKSRQSKYKRIIFSGYQSDIDTNRINDEL"
BASE COUNT 974 a 418 c 509 g 823 t
ORIGIN

alignment_scores:
Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A69697 ..
Align seg 1/1 to: A69697 from: 1 to: 2724

215 GlnSerlyVallysArgGlnIlePheSer 224
|||||
2644 CAATCTAAAGTTAAAGACAATATTTC 2673

seq_name: gb_pat:BD009884

seq_documentation_block:
LOCUS BD009884 2724 bp DNA linear PAT 31-JAN-2002
DEFINITION Recombinant toxin fragments.
ACCESSION BD009884
VERSION BD009884.1 GI:18638257
KEYWORDS JP 2001502890-A/8.
SOURCE unidentified.
ORGANISM unidentified.
unclassified.
1 (bases 1 to 2724)
REFERENCE 1 Shone, C.C., Quinn, C.P. and Foster, K.A.
AUTHORS Recombinant toxin fragments
TITLE Patent: JP 2001502890-A 8 06-MAR-2001;
JOURNAL MICROBIOLOGICAL RESEARCH AUTHORITY CAMR, THE SPEYWOOD LABORATORY LTD
OS Unidentified
COMMENT
PD JP 2001502890-A/8
PN 06-MAR-2001
PF 22-AUG-1997 JP 1998510524
PR 23-AUG-1996 GB 9617671.4.13-DEC-1996 GB 9625996.5 PI
CLIFFORD CHARLES SHONE, CONRAD PADRAIG QUINN, KEITH ALAN FOSTER PC

```

```

C12N15/31, C12N1/21, C12P21/02, C07K14/33, A61K38/16, A61K39/08 CC
Strandedness: Double:
CC Topology: Linear:
FH Key Location/Qualifiers
FT CDS 1..2724
Location/Qualifiers
1..2724
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 974 a 418 c 509 g 823 t
ORIGIN

alignment_scores:
Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BD009884 ..
Align seg 1/1 to: BD009884 from: 1 to: 2724

215 GlnSerlyVallysArgGlnIlePheSer 224
|||||
2644 CAATCTAAAGTTAAAGACAATATTTC 2673

seq_name: gb_htg:AC068893

seq_documentation_block:
LOCUS AC068893 154611 bp DNA linear HTG 30-AUG-2001
DEFINITION Homo sapiens chromosome 10 clone RP11-283C16, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
ACCESSION AC068893
VERSION AC068893.4 GI:13811856
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 154611)
REFERENCE 1 Smith, D.R.
AUTHORS Genome Therapeutics Corporation Sequencing Center: Human Genome
TITLE Sequence Data
JOURNAL Unpublished
2 (bases 1 to 154611)
REFERENCE 2 Smith, D.R.
AUTHORS Direct Submission
TITLE Submitted (11-MAY-2000) Genome Therapeutics Corporation, 100 Beaver
JOURNAL Street, Waltham, MA 02453, USA
COMMENT On Apr 27, 2001 this sequence version replaced gi:11545931.

----- Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
----- Project Information
Center project name: hg273
----- Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 167668 bases at least Q40
Consensus quality: 171341 bases at least Q30
Consensus quality: 173207 bases at least Q20
Insert size: 17782; sum-of-contigs
Quality coverage: 6.1x in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```


* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      1133: contig of 1133 bp in length
*      1233: gap of unknown length
*      1234      2305: contig of 1072 bp in length
*      2305      2405: gap of unknown length
*      2405      3675: contig of 1270 bp in length
*      3675      4875: gap of unknown length
*      4875      4975: contig of 1100 bp in length
*      4975      6456: gap of unknown length
*      6456      6557: gap of unknown length
*      6557      7888: contig of 1332 bp in length
*      7888      7989: gap of unknown length
*      7989      10074: contig of 2086 bp in length
*      10074     10175: gap of unknown length
*      10175     11247: contig of 1073 bp in length
*      11247     11348: gap of unknown length
*      11348     14041: contig of 2694 bp in length
*      14041     14142: gap of unknown length
*      14142     22538: contig of 8397 bp in length
*      22538     22638: gap of unknown length
*      22638     33993: contig of 11353 bp in length
*      33993     34094: gap of unknown length
*      34094     63953: contig of 29860 bp in length
*      63953     64054: gap of unknown length
*      64054     94115: contig of 30062 bp in length
*      94115     94216: gap of unknown length
*      94216     124020: contig of 29805 bp in length
*      124020     124121: gap of unknown length
*      124121     154611: contig of 30491 bp in length.
  
```

FEATURES
 source
 1. 154611
 Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone_lib="RPC1-11"
1. 1133
/note="assembly_name:Contig8"
misc_feature      1234..2305
/note="assembly_name:Contig11"
misc_feature      2406..3675
/note="assembly_name:Contig14"
misc_feature      3776..4875
/note="assembly_name:Contig18"
misc_feature      4976..6456
/note="assembly_name:Contig19"
misc_feature      6557..7888
/note="assembly_name:Contig20"
misc_feature      7989..10074
/note="assembly_name:Contig21"
misc_feature      10175..11247
/note="assembly_name:Contig22"
misc_feature      11348..14041
/note="assembly_name:Contig23"
misc_feature      14142..22538
/note="assembly_name:Contig24"
misc_feature      22639..33993
/note="assembly_name:Contig25"
misc_feature      34094..63953
/note="assembly_name:Contig26"
misc_feature      64054..94115
/note="assembly_name:Contig27"
misc_feature      94216..124020
/note="assembly_name:Contig28"
misc_feature      124121..154611
/note="assembly_name:Contig29"
BASE COUNT      45068 a 30855 c 31340 g 45938 t 1410 others
ORIGIN

```

alignment_scores:

Quality:	10.00	Length:	10
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x AC068893/rev ..

Align seg 1/1 to reverse of: AC068893 from: 1 to: 154611

62 ThrsEulerSerLeuArgSerAlaHisLeu 71

84838 ACTCTTGTGCTTACAGAGTGCACACCTC 84809

seq_name: gb_htg:AC021033

seq_documentation_block:

LOCUS AC021033

DEFINITION Homo sapiens chromosome 10 clone RP11-310A17, WORKING DRAFT

SEQUENCE, 24 unordered pieces.

ACCESSION AC021033.5 GI:13876476

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 166872)

AUTHORS Smith,D.R.

TITLE Sequence Data

JOURNAL Unpublished

2 (bases 1 to 166872)

REFERENCE Direct Submission

Submitted (12-JAN-2000) Genome Therapeutics Corporation, 100 Beaver

Street, Waltham, MA 02453, USA

On Apr 28, 2001 this sequence version replaced gi:9937742.

COMMENT

Genome Center

Center: Genome Therapeutics Corporation

Center code: GTC

Web site: http://www.genomecorp.com/

Contact: gtc-seqcenter@genomecorp.com

Project Information

Center project name: hg104

Summary Statistics

Sequencing vector: N/A

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 990315

Consensus quality: 181482 bases at least Q40

Consensus quality: 184094 bases at least Q30

Consensus quality: 185257 bases at least Q20

Insert size: 191465; sum-of-contigs

Quality coverage: 6.3x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1      1139: contig of 1139 bp in length
*      1140      1239: gap of unknown length
*      1239      1240      2469: gap of unknown length
*      2469      3971: contig of 1402 bp in length
*      3971      4071: gap of unknown length
*      4071      5550: contig of 1480 bp in length
*      5550      5651: gap of unknown length
*      5651      6874: contig of 1224 bp in length
  
```



```

* 6875 6974: gap of unknown length
* 6975 8613: contig of 1639 bp in length
* 8614 8713: gap of unknown length
* 8714 9769: contig of 1056 bp in length
* 9770 9869: gap of unknown length
* 9870 11095: contig of 1226 bp in length
* 11096 11195: gap of unknown length
* 11196 13378: contig of 2183 bp in length
* 13379 13478: gap of unknown length
* 13479 17449: contig of 3971 bp in length
* 17450 17549: gap of unknown length
* 17550 20361: contig of 2812 bp in length
* 20362 20462: gap of unknown length
* 20463 24773: contig of 4312 bp in length
* 24774 28965: gap of unknown length
* 28966 29065: gap of unknown length
* 29066 34516: contig of 5451 bp in length
* 34517 34616: gap of unknown length
* 34617 42471: contig of 7855 bp in length
* 42472 42571: gap of unknown length
* 42572 47944: contig of 5373 bp in length
* 47945 48044: gap of unknown length
* 48045 54711: contig of 6667 bp in length
* 54712 54811: gap of unknown length
* 54812 63169: contig of 8358 bp in length
* 63170 63268: gap of unknown length
* 63270 74889: contig of 11620 bp in length
* 74890 74990: gap of unknown length
* 74991 87514: contig of 12525 bp in length
* 87515 87614: gap of unknown length
* 87615 95455: contig of 7841 bp in length
* 95456 95555: gap of unknown length
* 95556 125405: contig of 29850 bp in length
* 125406 125505: gap of unknown length
* 125506 149661: contig of 24156 bp in length
* 149662 149762: gap of unknown length
* 149763 166872: contig of 17111 bp in length.

```

FEATURES

source

```

1..166872
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-310A17"
/clone_1bp="RPCT-11"
1..1139
/feature="assembly_name:Contig1"
1240..2468
/feature="assembly_name:Contig9"
2569..3970
/feature="assembly_name:Contig11"
4071..5350
/feature="assembly_name:Contig13"
5351..6874
/feature="assembly_name:Contig14"
6875..8613
/feature="assembly_name:Contig15"
8714..9769
/feature="assembly_name:Contig17"
9870..11095
/feature="assembly_name:Contig19"
11196..13378
/feature="assembly_name:Contig20"
13479..17449
/feature="assembly_name:Contig21"
17550..20361
/feature="assembly_name:Contig22"
20462..24773
/feature="assembly_name:Contig23"
24874..28965
/feature="assembly_name:Contig24"
29066..34516

```

```

/feature="assembly_name:Contig25"
clone_end:77"
34617..42471
/feature="assembly_name:Contig26"
42572..47944
/feature="assembly_name:Contig27"
48045..54711
/feature="assembly_name:Contig28"
54812..63169
/feature="assembly_name:Contig29"
63270..74889
/feature="assembly_name:Contig30"
74990..87514
/feature="assembly_name:Contig31"
87615..95455
/feature="assembly_name:Contig32"
95556..125405
/feature="assembly_name:Contig33"
125506..149661
/feature="assembly_name:Contig34"
149762..166872
/feature="assembly_name:Contig35"

```

```

BASE COUNT 47731 a 34417 c 33394 g 49026 t 2304 others
ORIGIN

```

```

alignment_scores:
Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

```

alignment_block:

```

```

US-09-528-682-1 x AC021033 ..

```

```

Align seg 1/1 to: AC021033 from: 1 to: 166872

```

```

62 ThrsEulerSerDeurGserAlaHlsleu 71
|||||
141955 ACTCTTTGTCTCCTTAGGAGTGCACACCTC 141984

```

```

seq_name: gb_pr:AC073367

```

```

seq_documentation_block:
LOCUS AC073367 168918 bp DNA linear PRI 25-OCT-2001
DEFINITION Homo sapiens chromosome 10 clone RP11-197M22, complete sequence.
ACCESSION AC073367
VERSION AC073367.10 GI:16418121
KEYWORDS HTG.
SOURCE human.

```

```

ORGANISM Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 168918)
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data

```

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 168918)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA

```

```

REFERENCE 3 (bases 1 to 168918)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-2001) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA

```

```

COMMENT On Oct 25, 2001 this sequence version replaced gi:15553107.
FEATURES
Location/Qualifiers
1..168918
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"

```



```

/clone="RP11-197M22"
BASE COUNT      49033 a 34070 c 34897 g 50918 t
ORIGIN

```

```

alignment_scores:
  Quality: 10.00      Length: 10
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x AC073367/rev ..

```

```

Align seg 1/1 to reverse of: AC073367 from: 1 to: 168918

```

```

62 TThSerLeuSerLeuArgSerAlaHisLeu 71
|||||
104393 ACTTCTTGTGCTAAGAGGTGACACCTCG 104364

```


/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF59601 +	7.00	107.44	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAT74890 -	7.00	102.46	2.3e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH07447 -	7.00	107.39	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAA62931 -	7.00	102.46	2.3e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH07353 -	7.00	107.36	1.2e+03	7	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA62681 -	7.00	102.39	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT56879 +	7.00	107.32	1.2e+03	77	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL04193 +	7.00	102.32	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV21397 +	7.00	107.32	1.2e+03	77	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF57366 -	7.00	102.25	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV63198 +	7.00	107.32	1.2e+03	77	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAF33756 -	7.00	102.23	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX39909 +	7.00	107.13	1.3e+03	80	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS71276 -	7.00	102.14	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX39907 +	7.00	107.02	1.3e+03	81	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS71372 -	7.00	102.14	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX39904 -	7.00	107.01	1.3e+03	81	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF94435 -	7.00	102.14	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA91656 -	7.00	107.01	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH82780 -	7.00	102.09	2.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA93596 -	7.00	106.98	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA82749 -	7.00	102.06	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA67186 +	7.00	106.98	1.3e+03	82	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF64031 -	7.00	102.06	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX39906 +	7.00	106.93	1.3e+03	82	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH84728 -	7.00	101.89	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV03029 +	7.00	106.81	1.3e+03	83	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ12229 +	7.00	101.86	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH06425 -	7.00	106.75	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAC91212 -	7.00	101.86	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS74361 -	7.00	106.75	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAS91196 +	7.00	101.84	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA93516 -	7.00	107.01	1.3e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ12230 +	7.00	101.83	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC54189 -	7.00	106.62	1.4e+03	85	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ12231 +	7.00	101.83	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC52617 -	7.00	106.60	1.4e+03	86	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ12231 +	7.00	101.83	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA13985 +	7.00	106.40	1.4e+03	88	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAF67124 +	7.00	101.80	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAA87642 +	7.00	106.38	1.4e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.AAF67124 +	7.00	101.79	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA167448 -	7.00	106.36	1.4e+03	8	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT44455 +	7.00	101.79	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA02641 -	7.00	106.22	1.4e+03	90	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59657 +	7.00	101.79	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS070364 -	7.00	106.19	1.4e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAO92524 -	7.00	101.76	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC38786 -	7.00	106.13	1.5e+03	91	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC54400 +	7.00	101.76	2.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS25976 +	7.00	106.11	1.5e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS68133 -	7.00	101.43	2.7e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAT50067 -	7.00	106.05	1.5e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAZ65344 +	7.00	101.30	2.7e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF14491 +	7.00	105.02	1.5e+03	93	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH84742 -	7.00	100.99	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS72626 -	7.00	105.63	1.6e+03	9	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF52125 +	7.00	100.97	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX20025 -	7.00	105.21	1.6e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS82129 +	7.00	100.96	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAO99460 -	7.00	105.09	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV20862 +	7.00	100.93	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT29627 -	7.00	104.95	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA46136 +	7.00	100.93	2.8e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.AAT01475 -	7.00	104.95	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA47422 -	7.00	100.85	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC35915 -	7.00	104.86	1.7e+03	10	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC38450 -	7.00	100.84	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS54152 +	7.00	104.83	1.7e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH84759 +	7.00	100.70	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH73931 -	7.00	104.83	1.7e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL07521 +	7.00	100.68	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH02888 -	7.00	104.83	1.7e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAO79705 -	7.00	100.63	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC47426 +	7.00	104.63	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAO79066 -	7.00	100.63	2.9e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC37980 +	7.00	104.58	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAS93383 -	7.00	100.63	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC37580 -	7.00	104.52	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAS94796 +	7.00	100.51	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC39323 -	7.00	104.45	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC47954 +	7.00	100.50	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA96224 -	7.00	104.45	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL26731 +	7.00	100.47	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS33667 -	7.00	104.44	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA33152 +	7.00	100.47	3.0e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX20824 -	7.00	104.44	1.8e+03	11	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAO79704 -	7.00	100.35	3.1e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS88441 +	7.00	104.43	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV09065 -	7.00	100.35	3.1e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77743 -	7.00	104.43	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX93382 -	7.00	100.35	3.1e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK77745 -	7.00	104.43	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAD26582 +	7.00	99.88	3.2e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAK78705 -	7.00	104.43	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.ABL24490 -	7.00	99.79	3.2e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS53102 +	7.00	104.43	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAO15170 -	7.00	99.60	3.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS78705 -	7.00	104.43	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAV43879 -	7.00	99.60	3.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS53102 +	7.00	104.43	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAO38091 -	7.00	99.52	3.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF6849 -	7.00	104.28	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS79387 -	7.00	99.40	3.4e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF6849 -	7.00	104.28	1.8e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS78706 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF67147 +	7.00	104.12	1.9e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX81526 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF64445 -	7.00	104.12	1.9e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX81887 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC33336 +	7.00	104.05	1.9e+03	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV15527 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS79299 +	7.00	103.69	2.0e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV16519 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS59328 -	7.00	103.65	2.0e+03	12	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX83888 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.AAT88180 -	7.00	103.65	2.0e+03	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC46027 -	7.00	99.35	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAF21703 +	7.00	103.39	2.1e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT44581 -	7.00	99.25	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAF67123 +	7.00	103.28	2.1e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD06040 +	7.00	99.19	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB12897 -	7.00	103.16	2.1e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS44736 -	7.00	99.11	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.AAZ06250 -	7.00	103.13	2.1e+03	13	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS43070 -	7.00	99.10	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAB14399 -	7.00	103.12	2.1e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS44908 -	7.00	99.05	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD07752 +	7.00	103.06	2.2e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AA12557 +	7.00	99.05	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.AAT26989 -	7.00	103.06	2.2e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAH16894 -	7.00	99.03	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAS958184 -	7.00	102.95	2.2e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAA35042 -	7.00	98.97	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS55315 +	7.00	102.95	2.2e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAT95233 -	7.00	98.97	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAC77954 -	7.00	102.71	2.3e+03	14	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.AAZ95234 -	7.00	98.97	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.AAS71553 -	7.00	102.59	2.3e+03	1	/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF99963 -	7.00	98.97	3.5e+03	7
/SIDS/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAF94170 -	7.00	102.56	2.3e+03	1					

/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA51961	7.00	90.54	1.1e+04	8	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE21613	7.00	68.66	1.7e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA51463	7.00	90.50	1.1e+04	8	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA51614	7.00	66.40	2.3e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL06304	7.00	89.50	1.2e+04	8	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH62688	7.00	64.57	2.9e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK86613	7.00	89.79	1.2e+04	9	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH65534	7.00	64.47	2.9e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK86614	7.00	89.34	1.3e+04	9	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE21544	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL02636	7.00	89.34	1.3e+04	9	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE21607	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:ABL02636	7.00	89.72	1.3e+04	9	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE21609	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA526706	7.00	88.65	1.4e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE21610	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:ABAT17191	7.00	88.64	1.4e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE21611	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:ABL32319	7.00	88.56	1.4e+04	10	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE21612	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAAL07031	7.00	88.39	1.4e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH14223	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK79984	7.00	88.39	1.4e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH14226	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI07030	7.00	88.39	1.4e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH69530	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK79983	7.00	88.39	1.4e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH65531	7.00	63.58	3.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA545478	7.00	88.33	1.4e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE22303	7.00	59.59	5.2e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:ABL33984	7.00	88.33	1.4e+04	11	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE22303	7.00	57.34	6.8e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA530105	7.00	88.12	1.5e+04	11	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH14489	7.00	57.34	6.8e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAO92323	7.00	88.00	1.5e+04	11	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAK91990	7.00	54.58	9.3e+05
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAK12985	7.00	87.88	1.5e+04	11	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH81490	7.00	53.17	1.0e+06
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA530100	7.00	87.79	1.5e+04	11	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAK95240	7.00	53.14	1.1e+06
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAAB1731	7.00	87.69	1.5e+04	12	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK96733	7.00	53.14	1.1e+06
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK56274	7.00	87.52	1.6e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI99683	7.00	45.45	2.2e+06
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AA559530	7.00	87.42	1.6e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI99682	7.00	45.44	2.2e+06
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAK13375	7.00	86.88	1.7e+04	13	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO32800	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL03288	7.00	86.88	1.7e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO48097	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO48062	7.00	86.82	1.7e+04	13	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO48098	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL15140	7.00	86.29	1.9e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO50101	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK58509	7.00	85.92	1.9e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV51698	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI05933	7.00	85.89	1.9e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAE58891	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI05934	7.00	85.89	1.9e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAE58892	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK73540	7.00	85.88	2.0e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAE58894	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK73537	7.00	85.88	2.0e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAE58950	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK73538	7.00	85.88	2.0e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAE79785	6.00	115.96	412.03
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK73538	7.00	85.37	2.0e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO21399	6.00	115.60	431.91
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL12468	7.00	85.37	2.0e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO23258	6.00	115.60	431.91
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL18210	7.00	85.37	2.1e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO23258	6.00	115.25	451.70
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAK20510	7.00	85.08	2.2e+04	17	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO32903	6.00	115.25	451.70
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI19449	7.00	84.89	2.2e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO48101	6.00	115.25	451.70
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAI56114	7.00	84.86	2.2e+04	17	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAE58935	6.00	115.25	451.70
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ42200	7.00	84.86	2.2e+04	17	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAE58935	6.00	115.25	451.70
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK77679	7.00	84.85	2.2e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO51010	6.00	114.92	471.41
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAK66943	7.00	84.21	2.4e+04	1	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV53725	6.00	114.92	471.41
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL16796	7.00	83.88	2.5e+04	20	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV47548	6.00	114.92	471.41
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAK13229	7.00	83.78	2.6e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV16288	6.00	114.92	471.41
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL04804	7.00	83.65	2.6e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAI12729	6.00	114.92	471.41
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL19002	7.00	83.24	2.7e+04	22	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE58546	6.00	114.92	471.41
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAH11658	7.00	83.24	2.7e+04	22	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO24238	6.00	114.92	471.41
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAH05882	7.00	82.77	2.7e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO24238	6.00	114.29	510.63
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:ABAT18618	7.00	82.66	3.0e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAO46053	6.00	114.00	530.14
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABAT19048	7.00	82.66	3.0e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAE21258	6.00	114.00	530.14
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAH59608	7.00	82.25	3.1e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAZ20286	6.00	114.00	530.14
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH11177	7.00	81.91	3.3e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH38459	6.00	113.72	549.58
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK70447	7.00	81.53	3.4e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAO43025	6.00	113.72	549.58
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK1264	7.00	81.28	3.5e+04	2	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAE51319	6.00	113.72	549.58
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA532516	7.00	80.63	3.7e+04	3	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH65501	6.00	113.72	549.58
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABL13678	7.00	80.60	3.7e+04	3	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAH65501	6.00	113.45	568.96
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:ABL12964	7.00	80.54	3.9e+04	32	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE03367	6.00	113.45	568.96
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAK12954	7.00	80.49	3.9e+04	33	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAE03367	6.00	113.45	568.96
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAK1294	7.00	79.49	4.4e+04	37	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAH3173	6.00	113.19	586.28
/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAK06831	7.00	78.38	4.9e+04	37	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAE72905	6.00	112.94	607.55
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK61318	7.00	78.34	5.1e+04	44	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAK09492	6.00	112.94	607.55
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ61522	7.00	78.29	5.1e+04	44	/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK33782	6.00	112.69	626.76
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAK75080	7.00	77.97	5.3e+04	44	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO20875	6.00	112.69	626.76
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:ABAT16830	7.00	76.80	6.2e+04	5	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAO51639	6.00	112.69	626.76
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAK67426	7.00	76.35	6.6e+04	5	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAO63928	6.00	112.46	684.10
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAK56471	7.00	76.01	6.9e+04	58	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAE90673	6.00	112.46	684.10
/SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAK63782	7.00	76.01	6.9e+04	6	/SID5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAO24243	6.00	112.01	684.10

/SID55_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAV7592-	6.00	112.01	684.10	3	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAK57505-	6.00	103.52	2.0e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAF82949+	6.00	111.80	703.11	34	/SID5_gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAAC67158-	6.00	103.54	2.1e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF8283+	6.00	111.80	703.11	34	/SID5_gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAAC14718+	6.00	103.52	2.1e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAX27561+	6.00	111.59	722.02	35	/SID5_gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAAF69389+	6.00	102.97	2.2e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAAO99697+	6.00	110.63	816.29	40	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS49347+	6.00	102.77	2.2e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1995.DAT:AAO975657+	6.00	110.63	816.29	40	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAS49362+	6.00	102.77	2.2e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAAT26218+	6.00	110.63	816.29	40	/SID5_gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAS66449+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAF65126+	6.00	110.63	816.29	40	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT77747+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAF51120+	6.00	110.63	816.29	40	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT77748+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAZ48342+	6.00	110.63	816.29	40	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT77751+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAZ48342+	6.00	110.63	816.29	40	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT77752+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAZ484802+	6.00	110.63	816.29	40	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT77755+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAAK68483+	6.00	110.46	835.01	42	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT77756+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAAO36760+	6.00	110.29	853.69	41	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAB81147+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK78294+	6.00	109.79	909.53	45	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAB81147+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK78295+	6.00	109.79	909.53	45	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAB81147+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAZ67727+	6.00	109.48	946.59	47	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAB81147+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAZ698181-	6.00	109.48	946.59	47	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAB81147+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAZ698181-	6.00	109.48	946.59	47	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAB81147+	6.00	102.71	2.3e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAAT484801+	6.00	109.48	946.59	47	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAT484801+	6.00	102.31	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAAZ484801+	6.00	109.04	1.0e+03	50	/SID5_gcgdata/geneseq/geneseq-emb1/NA1998.DAT:AAK12311-	6.00	102.25	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK40202-	6.00	109.04	1.0e+03	50	/SID5_gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAK16317+	6.00	102.25	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34267-	6.00	109.04	1.0e+03	50	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34397-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	109.04	1.0e+03	50	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAF76756+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAAL33378+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH87859-	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAL34268-	6.00	102.19	2.4e+03	1
/SID55_gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAAH89395+	6.00	108.89	1.0e+03	51	/SID5_gcgdata/geneseq/geneseq-emb1/NA2001A				

/SID5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.AA06058 +	6.00	99.98	3.2e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK24713 -	6.00	98.43	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC0159 +	6.00	99.94	3.2e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK50713 -	6.00	98.43	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC02706 +	6.00	99.90	3.2e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK127726 -	6.00	98.43	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC15089 +	6.00	99.90	3.2e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.AAK56693 -	6.00	98.43	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF68390 -	6.00	99.90	3.2e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.AAK00041 -	6.00	98.40	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS50048 -	6.00	99.86	3.2e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAK55287 -	6.00	98.40	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS50048 +	6.00	99.82	3.3e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAK55370 -	6.00	98.40	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK40062 +	6.00	99.63	3.3e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK54176 +	6.00	98.40	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK24001 +	6.00	99.63	3.3e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF85499 -	6.00	98.40	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK52351 -	6.00	99.63	3.3e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF17746 +	6.00	98.40	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC55231 +	6.00	99.48	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAK12385 -	6.00	98.35	3.9e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC05630 -	6.00	99.44	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.AAK35353 +	6.00	98.30	4.0e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA31618 +	6.00	99.44	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAF54584 +	6.00	98.30	4.0e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAA17115 -	6.00	99.44	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAF75489 +	6.00	98.27	4.0e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK47095 +	6.00	99.40	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS49278 +	6.00	98.24	4.0e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB64977 +	6.00	99.40	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.AAS50420 +	6.00	98.24	4.0e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK39140 +	6.00	99.40	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAV75328 +	6.00	98.20	4.0e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK119949 +	6.00	99.40	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAK14438 -	6.00	98.17	4.0e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI145147 +	6.00	99.40	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB76445 +	6.00	98.08	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI105617 +	6.00	99.40	3.4e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK25080 +	6.00	98.08	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.AAT15268 -	6.00	99.37	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK51087 +	6.00	98.08	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC44265 -	6.00	99.37	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI157125 +	6.00	98.08	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC19356 +	6.00	99.29	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAH72234 -	6.00	98.08	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAA478916 +	6.00	99.25	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB74233 -	6.00	98.05	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA45191 +	6.00	99.22	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAA39218 -	6.00	98.05	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC40380 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAL19758 +	6.00	98.05	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA44324 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK22694 -	6.00	98.05	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA51313 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK48861 -	6.00	98.05	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA69312 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI12634 -	6.00	98.05	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK43414 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI16460 -	6.00	98.05	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI124193 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.AAA07258 +	6.00	99.02	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI149479 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAA70055 +	6.00	98.02	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI109758 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK18265 +	6.00	98.02	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.ABI122203 +	6.00	99.18	3.5e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK44163 +	6.00	98.02	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABI74035 +	6.00	99.15	3.6e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI50170 +	6.00	98.02	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA39094 +	6.00	99.15	3.6e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF93105 +	6.00	98.02	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK22489 +	6.00	99.15	3.6e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF94536 +	6.00	98.02	4.1e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK48691 +	6.00	99.15	3.6e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC52642 +	6.00	97.83	4.2e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI126527 +	6.00	99.15	3.6e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA31101 +	6.00	97.77	4.2e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI154485 +	6.00	99.15	3.6e+03	17	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA31183 -	6.00	97.77	4.2e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.AAH86238 -	6.00	99.11	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAS3890 +	6.00	97.75	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.AAH86493 +	6.00	99.07	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC03561 -	6.00	97.72	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC01785 -	6.00	99.07	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC19668 -	6.00	97.69	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.AAC071823 +	6.00	99.04	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC11607 -	6.00	97.66	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC94212 +	6.00	99.04	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA75422 +	6.00	97.66	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA47801 +	6.00	99.00	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK23087 +	6.00	97.66	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA65591 +	6.00	99.00	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK50053 +	6.00	97.66	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.ABA32776 +	6.00	99.00	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI12741 +	6.00	97.66	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK14095 +	6.00	99.00	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI155939 +	6.00	97.66	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK39838 +	6.00	99.00	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK59368 -	6.00	97.60	4.3e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI120648 +	6.00	99.00	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.AAO59356 +	6.00	97.57	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI145858 +	6.00	99.00	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.AAO83975 +	6.00	97.57	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI106748 -	6.00	99.00	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAO30121 -	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.AAAB10782 +	6.00	98.97	3.6e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.AAG76781 -	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC434362 -	6.00	98.97	3.7e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAA55111 +	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.AAA111759 -	6.00	98.86	3.7e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.AAV54751 -	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC10004 -	6.00	98.83	3.7e+03	20	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.AAA48202 -	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.AAO72529 -	6.00	98.69	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.AAV36154 -	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.AAT39206 +	6.00	98.66	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAA90378 -	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.AAT39211 +	6.00	98.66	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB849158 +	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.AA053936 +	6.00	98.66	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAA60904 +	6.00	97.54	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.AA054054 +	6.00	98.66	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK78852 +	6.00	97.51	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAAS0481 +	6.00	98.66	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAH82016 +	6.00	97.51	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAB72364 -	6.00	98.59	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.AAV77424 -	6.00	97.48	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAA38301 -	6.00	98.59	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.AAA89260 +	6.00	97.48	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK20981 +	6.00	98.59	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF93605 +	6.00	97.48	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK47143 -	6.00	98.59	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.AAA067732 +	6.00	97.46	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI125658 -	6.00	98.59	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK73924 -	6.00	97.46	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI152878 -	6.00	98.59	3.8e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK22377 -	6.00	97.46	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.AAC53936 +	6.00	98.46	3.9e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAK49544 -	6.00	97.46	4.4e+	21
/SID5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.AAV290942 +	6.00	98.46	3.9e+03	21	/SID5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAI154374 -	6.00	97.46	4.4e+	21


```

XX OS Escherichia coli.
XX PN JP04079898-A.
XX PD 13-MAR-1992.
XX PF 23-JUL-1990; 90JP-0194208.
XX PR 23-JUL-1990; 90JP-0194208.
XX PA (KITA ) KITASATO RES INST.
XX DR WPI; 1992-137930/17.
XX PT DNA and RNA probe - for simultaneous detection of Vibrio cholerae
XX PF and entero-toxicogenic E. coli
XX PS
XX PS Claim 1; Page 1; 15pp; Japanese.
XX CC The thermostable toxin (LTb) gene sequence is derived from
XX CC enterotoxigenic E. coli. An LTh-ST1a-ST1b probe, wherein all
XX CC all or part of the sequences of the three toxin genes represented
XX CC in AA023864-66 are ligated to each other by T4DNA ligase, may be used
XX CC for the detection of Vibrio cholerae and enterotoxigenic E.coli
XX CC simultaneously in a simple way.
XX CC The method has the same sensitivity as the dangerous radio-isotope
XX CC detection method.
XX SO Sequence 1148 BP; 399 A; 205 C; 229 G; 315 T; 0 other:

```

```

alignment_scores:
    Quality: 208.00      Length: 208
    Ratio: 1.000         Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x AA023864 ..

```

```

Align seg 1/1 to: AA023864 from: 1 to: 1148

```

```

5 LeuTyrArgAlaAspSerArgProAspGluIleLysArgSerGly 21
|||||
67 TTATACCGTGGCTGACTGTAGACCCCGACATGAATAAACGTTCCGAGG 116
|||||
21 yLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThiMet 38
|||||
117 TCTTATGCCAGAGGCGCATATGAGTACTCGATAGAGAACTCAATGA 166
|||||
38 snIleAsnLeuTyrAspHisAlaArgGlyThiGlnThrGlyPheVal 54
|||||
167 ATATTATCTTTATGTATCATCCGAGAGGAAACACCAACCGGCTTGTCA 216
|||||
55 TyrAspAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHis 71
|||||
217 TATGATGACGAGATATGTTCACTCTCTTACGTTGAGAAAGCTCCTCA 266
|||||
71 uAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrVal 88
|||||
267 AGCAGGACAGTCTATTTATCAGAGATATCCACTTACTATATATATGTT 316
|||||
88 leAlaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTyr 104
|||||
317 TAGCGACAGACCAAAATATGTTTAAATGTTATATGATATAGCGGTATAC 366
|||||
105 SerProHisProTyrGluGlnGluValSerAlaLeuGlyGlyIleProTyr 121
|||||
367 AGCCCTCACCAATATGACAGAGGTTTCGCGTTAGTGGTGAATACATA 416
|||||
121 rSerGlnIleTyrGlyTyrTyrArgValAsnPheGlyValIleAspGlu 138
|||||
417 TTCTCAGATATATGATGATGATGTTTATTTTGGTGTGATGATGATGAC 466

```

```

138 rLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
|||||
467 GATTACATCTGTTACAGAGGAATATAGAACCGGTATTTACAGAAATCTGAAT 516
|||||
155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHis 171
|||||
517 ATAGCTCCGGCAGAGAGATGTTACAGATTATGACAGGTTTCCACCGGATCA 566
|||||
171 sGlnAlaTTPArgGluGluProTyrPheHisAlaProGlnGlyCysG 188
|||||
567 CCAAGCTTGAGAGAGAAACCTGATTCATCATGACCCCAAGGTTGTG 616
|||||
188 lYAsnSerSerArgThrIleThrGlyAspPheCysAsnGluGluThrGln 204
|||||
617 GAAATTCATCAGACACATTTACAGGTGATCTTGTATATGAGAGACCCAG 666
|||||
205 AsnLeuSerThrIleTyrLeuArg 212
|||||
667 AATCTGACGACATATATATCTCAGG 690
|||||
seq_name: /SID5/gc9data/geneseq/geneseqn-emb1/NA2000.DAT:AAA51106
seq_documentation_block:
ID AAA51106 standard; cDNA; 782 BP.
XX
XX AAA51106;
XX
XX 26-SEP-2000 (first entry)
XX
XX DE Plant-optimized E. coli heat labile toxin A subunit coding sequence.
XX
XX KM Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
XX KM adjuvant; anti-bacterial; ss.
XX OS Escherichia coli.
XX
XX Key Location/Qualifiers
XX CDS 3..782
XX FT /*tag= a
XX FT /product= heat-labile_toxin-A
XX FT sig_peptide 3..56
XX FT /*tag= b
XX FT mat_peptide 57..779
XX FT /*tag= c
XX
XX WO200037609-A2.
XX
XX PD 29-JUN-2000.
XX
XX PF 22-DEC-1999; 99MO-US30747.
XX
XX PR 22-DEC-1998; 98US-0113507.
XX
XX PA (BOYCE-) BOYCE THOMPSON INST PLANT RES.
XX PA (MASO/) MASON H S.
XX PA (ARNT/) ARNTZEN C J.
XX
XX PI Mason HS, Arntzen CJ;
XX
XX DR WPI; 2000-442653/38.
XX DR P-PSDB; AA96646.
XX
XX PT New polynucleotides encoding LT-A or CT-A polypeptides for the
XX PT transformation of plant cells, useful in immunogenic compositions to
XX PT elicit immune responses in animals
XX
XX PS Example 1; Fig 1; 103pp; English.
XX
XX This plant-codon optimized cDNA encodes a synthetic Escherichia coli
XX CC heat-labile toxin (LT) A subunit (LT-A). The sequence contains
XX CC plant-preferred codons and eliminates sequence motifs associated with

```


CC spurious mRNA processing. A single codon insertion (GTG encoding valine)
 CC was made to accommodate the creation of a NcoI restriction site around the
 CC initiator methionine codon. Novel polynucleotides encode a mutant LT-A
 CC polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit
 CC (CT-A) polypeptide, which have reduced enzyme activity as compared to the
 CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
 CC is altered to a plant preferred codon. The polynucleotide further
 CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
 CC subunit (CT-B). The polynucleotides are useful for the transformation of
 CC plant cells for the production of transgenic plants to produce edible
 CC vaccines, especially oral vaccines in transgenic plants for the
 CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The
 CC mutant polypeptides are also useful as adjuvants.

XX Sequence 782 BP; 225 A; 186 C; 191 G; 180 T; 0 other;

alignment_scores:
 Quality: 184.00 Length: 184
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA51106 ..

Align seg 1/1 to: AA51106 from: 1 to: 782

5 LeuTYrARgAlaAspSerArgProProAspGluIleLysArgSerGly 21
 |||||
 72 CTCTATAGGGCTGACCTTAGGCCCCAGATGATCAAGAGCTGTGAGG 121
 |||||
 21 yLeuMetProARgGlyHISaNGluTYrPheAspArgGlyTHrGluMetA 38
 |||||
 122 TCTCATGCGCCAAGGGGACACAAATGACTTTGATAGGGGAACCAATGA 171
 |||||
 38 snLLeuSnLeuTYrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
 |||||
 172 ACATCAACCTCTATGACCATGTAGGGGAACCAACTGGCTTTGTAGG 221
 |||||
 55 TYrAspAspGlyTYrValSerThrSerLeuSerLeuArgSerAlaHisLe 71
 |||||
 222 TATGATGATGATATGTCCTCACCCTCCTTAGCTGAGGCTGCTCCTT 271
 |||||
 71 uAlaGlyGlnSerIleLeuSerGlyTYrSerThrTYrTYrIleTYrValI 88
 |||||
 272 GGCAGGACATFCATCCTCTCAGGATATCCACCTACTACATCATGTGA 321
 |||||
 88 lElaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTYr 104
 |||||
 322 TTGCTACAGACCAACATGTCTCAATGTGAATGTGTGGAGAGTGTAT 371
 |||||
 105 SerProHisProTYrGlnGlnGluValSerAlaLeuGlyGlyIleProTY 121
 |||||
 372 AGGCTTACCCATATGAGCAAGAGTGTCTGCTTTGGGTGGAATCCATA 421
 |||||
 121 rSerGlnIleTYrGlyTYrTYrArgValAsnPheGlyValIleAspGluA 138
 |||||
 422 CTCCTCAATCTATGATGTAGTGGGAGACTTTGGGTGATGATGAGA 471
 |||||
 138 rGluLeuHisAspAsnArgGluTYrArgAspArgTYrTYrArgAsnLeuAsn 154
 |||||
 472 GGCCTCCATAGGAATAGGAGTATAGGACAGGTACTATAGAACTCTAAC 521
 |||||
 155 lAlaAlaProAlaGluAspGlyTYrArgLeuAlaGlyPheProProAspH 171
 |||||
 522 ATAGCTCCACAGAGAGATGCTTAGTGTGCAAGTTTCCACACAGACCA 571
 |||||
 171 sGlnAlaTrpArgGlnGluProTrpIleHisHisAlaProGlnGlyCysG 188
 |||||
 572 CCAAGCTGTGAGGAGAGAGCCCTGATCCACATGCACCAAGGTTGTG 621
 |||||
 188 Ty 188
 ||

622 GA 623

seq_name: /SID5/gcdata/geneseq/geneseqn-emb1/NA2000.DAT:AA51544

seq_documentation_block:

ID AA51544 standard; cDNA: 782 BP.

XX AA51544;

XX 26-SEP-2000 (first entry)

XX Plant-optimized E. coli LT-A G192 mutant coding sequence.

XX Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;

XX adjuvant; anti-bacterial; R192G; ss.

XX Escherichia coli.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 3..782

FT /tag= a

FT sig_peptide 3..59

FT /tag= b

FT mat_peptide 60..779

FT /tag= c

PN W0200037609-A2.

XX 29-JUN-2000.

XX 22-DEC-1999; 99WO-US30747.

XX 22-DEC-1998; 98US-0113507.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES.

XX (MASO-) MASON H S.

XX (ARNT/) ARNTZEN C J.

XX Mason HS, Arntzen CJ;

XX WPI: 2000-442653/38.

XX P-PSDB: AAY96648.

XX New polynucleotides encoding LT-A or CT-A polypeptides for the

XX transformation of plant cells, useful in immunogenic compositions to

XX elicit immune responses in animals

XX Example 3; Page -: 103pp; English.

XX This cDNA encodes a mutant R192G Escherichia coli heat-labile toxin (LT)

XX A subunit (LT-A). The nucleotide at position 631 was changed from the

XX wild-type A to G which causes a change of arginine to glycine in the

XX mature protein at residue 192. The sequence contains plant-preferred

XX codons and eliminates sequence motifs associated with spurious mRNA

XX processing. A single codon insertion (GTG encoding valine) was made to

XX accommodate the creation of a NcoI restriction site around the

XX initiator methionine codon. Novel polynucleotides encode a mutant LT-A

XX polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit

XX (CT-A) polypeptide, which have reduced enzyme activity as compared to

XX the wild-type LT-A or CT-A polypeptide and where at least one of the

XX codons is altered to a plant preferred codon. The polynucleotide further

XX comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B

XX subunit (CT-B). The polynucleotides are useful for the transformation of

XX plant cells for the production of transgenic plants to produce edible

XX vaccines, especially oral vaccines in transgenic plants for the

XX prophylactic or therapeutic treatment against E. coli or V. cholerae.

XX The mutant polypeptides are also useful as adjuvants.

XX Note: This sequence does not appear in the specification. It was

XX constructed from the wild type LT-A cDNA shown in AA51106 which is

XX given in Figure 1 of the specification.

sq Sequence 782 BP; 224 A; 186 C; 192 G; 180 T; 0 other;

alignment_scores:

Quality: 184.00 Length: 184
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AAN51544 ..

Align seg 1/1 to: AAN51544 from: 1 to: 782

```

5 LeuTYrARgAlaAspSerArgProProAspGluIleLysArgSerGIYL 21
72 CTCTATAGGCTGACTAGCCGCCAGATGATGATCAAGAGGTCTGGAGG 121
21 yLeuMetProARgGlyhIsAnGluTYrPheAspARgLYThrgInMeTA 38
122 TCTCATGCCAAGGGGACACATAGTACTTTGATAGGGGAACCAATGA 171
38 snIleAsnLeuTYrAspHisAlaArgGlyThrgInThrgIlyPheValArg 54
172 ACATCAACCTCTATGACCCTAGGGGAACCCAACTGGCTTTGTGAGG 221
55 TYrAspAspGIYTYrValSerThrSerLeuSerLeuArgSerAlaHisIle 71
222 TATGATATGATATGATGTGTCCACTCCCTTAGCTAGAGGTCTGCTACT 271
71 uAlaGlyGlnSerIleLeuSerGIYTYrSerThrTYrTYrIleTYrValI 88
272 GGCGAGGACATTCATCCTCTCAGGATCTCCACTACTACTATCATGTGA 321
88 leaIaThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTYr 104
322 TTGGCTACAGACCAACATGTCTCATGTGATGTGTGGGAGTGTAT 371
105 SerProHisProTYrGIUGlnGluValSerAlaLeuGlyIleProTY 121
372 AGCCCTCACCCTATGAGCAAGAGGTCTGCTTGGGTGGAAATCCCAT 421
121 rSerGlnIleTYrGlyTYrTYrArgValAsnPheGlyValIleAspGlu 138
422 CTCCTCAATCTATGATGATGATAGGTGACTTGTGTGATGATGATGA 471
138 rGleuHisArgAsnArgGluTYrArgAspArgTYrTYrArgAsnLeuAsn 154
472 GGCTCATAGGAATAGGAGATATAGGACAGGTCTATAGCAACCTCAAC 521
155 IleAlaProAlaGluAspGlyTYrArgLeuAlaGlyPheProProAspHI 171
522 ATAGCTTCACGACAGAGATGCTTATAGGTGGCAGGTTTCCACACAGACA 571
171 sGlnAlaTYrPARgGluGluProTYrIleHisHisAlaProGlnIlycysG 188
572 CCAAGCCTGGAGGAGAGGCCCTGATTCACCATGACACACAGAGTTGTG 621
188 y 188
622 GA 623
seq_name: /STDS/gcgdata/geneseq/geneseqn-emb1/NM1985.DAT: AAN50206
seq_documentation_block:
ID AAN50206 standard; DNA; 777 BP.
XX
AC AAN50206;
XX
DT 30-OCT-1991 (first entry)
XX
DE Sequence of the pig scours heat labile toxin (LT) LTA gene of the
XX site directed mutant SDM1.
XX
```

KW Pig scours vaccine; toxin; diarrhoea; ss.

XX E.coli NCIB 11932.

OS Key Location/Qualifiers

FT CDS 1..777

FT misc_difference 235..237

FT /tag= a

FT /tag= b

FT /note= "TCC in native sq"

XX EP145486-A.

XX 19-JUN-1985.

XX 12-DEC-1984; 84EP-0308620.

XX 12-DEC-1983; 83GB-0033131.

XX (GLAX) GLAXO GROUP LTD.

XX Hayes MV, Harford S, Ross GW;

XX WPI; 1985-148358/25.

XX P-PsDB; AAP50191.

XX New toxoid as inactivated form of toxin for use in vaccines - 1s

XX obd. from organism transformed by gene

XX Example; Fig 2; 61p; English.

XX AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of

XX the site directed mutant SDM1 (see AAN50206) is inactive. The

XX inventors claim a vaccine prepn. active against pig scours which

XX contains an inactivated LTA component, together with additional K88

XX antigens opt. with whole cells comprising the antigens or contg. the

XX inactivated LTA.

XX Sequence 777 BP; 255 A; 135 C; 164 G; 223 T; 0 other;

alignment_scores:

Quality: 179.00 Length: 179
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AAN50206 ..

Align seg 1/1 to: AAN50206 from: 1 to: 777

```

62 ThrSerLeuSerLeuArgSerAlaHisIleuAlaGlyInsIleLeuSe 78
238 ACTCTCTTACTTGTGAGAAAGTCTCAGCTTACAGACAGTCTAATATATC 287
78 rGlyTYrSerThrTYrTYrIleTYrValIleAlaThrAlaProAsnMetP 95
288 AGGATATTCCACTTACTATATATATATAGTATAGCGACACCAATATATG 337
95 heAsnValAsnAspValLeuGlyValTYrSerProHisProTYrGIUGln 111
338 TTATATGTTAATGATGATATAGCGGATATACAGCCCTCACCCATATGACAG 387
112 GluValSerAlaLeuGlyGlyIleProTYrSerGlnIleTYrGlyTYrTY 128
388 GAGCTTCTCGCTTAGGATGAGTACCAATATCTCAGATATGTGATGATGTA 437
128 rArgValAsnPheGlyValIleAspGluArgLeuHisArgAsnArgGluT 145
438 TCGTGTATATTTGTGTGATGATGAGAACGATTCATCGTGAACAGGAGAT 487
145 yArgAspArgTYrTYrArgAsnLeuAsnIleAlaProAlaGluAspGly 161
|||||
```


PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
 PA (MASO/) MASON H S.
 PA (ARNT/) ARNTZEN C J.
 XX
 PI Mason HS, Arntzen CJ;
 XX
 DR WPI; 2000-442653/38.
 DR P-PSDB; AAY96647.
 XX
 PT New polynucleotides encoding LT-A or CT-A polypeptides for the
 PT transformation of plant cells, useful in immunogenic compositions to
 PT elicit immune responses in animals
 PS
 PS Example 2; Page -: 103pp; English.
 XX
 CC This cDNA encodes a mutant S63k Escherichia coli heat-labile toxin (LT)
 CC A subunit (LT-A). The codon at nucleotide positions 246-248 was changed
 CC from the wild-type TCC to AAG which causes a change of serine to lysine
 CC in the mature protein at residue 63. The sequence contains
 CC plant-preferred codons and eliminates sequence motifs associated with
 CC spurious mRNA processing. A single codon insertion (GTG encoding valine)
 CC was made to accommodate the creation of a NcoI restriction site around
 CC the initiator methionine codon. Novel polynucleotides encode a mutant
 CC LT-A polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A
 CC subunit (CT-A) polypeptide, which have reduced enzyme activity as
 CC compared to the wild-type LT-A or CT-A polypeptide and where at least one
 CC of the codons is altered to a plant preferred codon. The polynucleotide
 CC further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or
 CC a CT B subunit (CT-B). The polynucleotides are useful for the
 CC transformation of plant cells for the production of transgenic plants to
 CC produce edible vaccines, especially oral vaccines in transgenic plants
 CC for the prophylactic or therapeutic treatment against E. coli or V.
 CC Note: This sequence does not appear in the specification. It was
 CC constructed from the wild type LT-A cDNA shown in AAA51106 which is
 CC given in Figure 1 of the specification.
 XX
 SQ Sequence 782 BP; 227 A; 184 C; 192 G; 179 T; 0 other;

alignment_scores:
 Quality: 125.00 Length: 125
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AAA51147 ...

Align seg 1/1 to: AAA51147 from: 1 to: 782

64 LeuSerLeuArgSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTy 80
 |||||||
 249 CTTCAGCTTGTAGCTGCTCCTCCTGCGAGCAATCCATCCCTCAGGATA 298
 |||||||
 80 rSerThrTyTrIleTyValIleAlaThrAlaProAsnMetPheAsnV 97
 |||||||
 299 CTCACCTACTACATCTATGTGATTGTCACGCCAACCAACCTGTCATG 348
 |||||||
 97 alaAsnValIleuGlyValTySerProHisProTyGlnGlnVal 113
 |||||||
 349 TGATGATGTGTTGGAGTGTATAGCCCTCACCATTATGACAGAGCTG 398
 |||||||
 114 SerAlaLeuGlyGlyIleProTySerGlnIleTyGlyTyTyArgVa 130
 |||||||
 399 TCTGCTTGGGTGGATCCATCTCCCAAACTATGATGATGATAGGT 448
 |||||||
 130 IAsnPhcGlyValIleAspGluArgLeuHisArgAsnArgGlnTyArgA 147
 |||||||
 449 GAACCTTTGGTGTGATGTAGAGAGCTCCATAGGAATAGGAGATAGGG 498
 |||||||
 147 SPATGTYTYTYArgAsnLeuAsnIleAlaProAlaGluAspGlyTYArg 163
 |||||||
 499 ACAGGTCTATAGAACCTCAACATAGCTCCAGCAGAGATGTTATAGG 548

164 LeuAlaGlyPheProAspHisGlnAlaTrpArgGlnGluProTyrPI 180
 |||||||
 549 TTGGCAGGTTTCCACACCAACCAAGCCTGAGGAGGAGAGCCCTGGAT 598
 |||||||
 180 eHisHisAlaProGlnGlyGly 188
 |||||||
 599 CCACCATGCACACACAGTGTGTGA 623

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51545

seq_documentation_block:
 ID AAA51545 standard; cDNA; 782 BP.
 XX
 AC AAA51545;
 XX
 DT 26-SEP-2000 (first entry)
 XX
 DE Plant-optimized E. coli LT-A R72 mutant coding sequence.
 XX
 KW Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
 KM adjuvant; anti-bacterial; A72R; ss.
 XX
 OS Escherichia coli.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..782
 FT /tag= a
 FT /product= heat-labile_toxin-A
 FT sig_peptide 3..59
 FT /tag= b
 FT mat_peptide 60..779
 FT /tag= c
 XX
 PN WO200037609-A2.
 XX
 PD 29-JUN-2000.
 XX
 XX 22-DEC-1999; 99WO-US30747.
 PF
 XX 22-DEC-1998; 98US-0113507.
 PR
 XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
 PA (MASO/) MASON H S.
 PA (ARNT/) ARNTZEN C J.
 XX
 PI Mason HS, Arntzen CJ;
 XX
 DR WPI; 2000-442653/38.
 DR P-PSDB; AAY96649.
 XX
 PT New polynucleotides encoding LT-A or CT-A polypeptides for the
 PT transformation of plant cells, useful in immunogenic compositions to
 PT elicit immune responses in animals
 PS
 PS Example 4; Page -: 103pp; English.
 XX
 CC This cDNA encodes a mutant A72R Escherichia coli heat-labile toxin (LT)
 CC A subunit (LT-A). The nucleotides at position 273-274 were changed from
 CC the wild-type GC to AG which causes a change of alanine to arginine in
 CC the mature protein at residue 72. The sequence contains plant-preferred
 CC codons and eliminates sequence motifs associated with spurious mRNA
 CC processing. A single codon insertion (GTG encoding valine) was made to
 CC accommodate the creation of a NcoI restriction site around the
 CC initiator methionine codon. Novel polynucleotides encode a mutant LT-A
 CC polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit
 CC (CT-A) polypeptide, which have reduced enzyme activity as compared to
 CC the wild-type LT-A or CT-A polypeptide and where at least one of the
 CC codons is altered to a plant preferred codon. The polynucleotide further
 CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
 CC subunit (CT-B). The polynucleotides are useful for the transformation of
 CC plant cells for the production of transgenic plants to produce edible

CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae.
CC The mutant polypeptides are also useful as adjuvants.
CC Note: This sequence does not appear in the specification. It was
CC constructed from the wild type LT-A cDNA shown in AA051106 which is
CC given in Figure 1 of the specification.

XX Sequence 782 BP; 226 A; 185 C; 191 G; 180 T; 0 other;

alignment_scores:

Quality: 116.00 Length: 116
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA051545 ..

Align seg 1/1 to: AA051545 from: 1 to: 782

73 GlnGlnSerIleLeuSerGlyTyrSerThrTyrTrpIleValIleAl 89
|||||
276 GGACAAATCCATCTCTCAGGATCTCCACCTACTACATCTATGTGATTGC 325
89 aThrAlaProAsnMetPheAsnValAsnAspValLeuGlyValTyrSerP 106
|||||
326 TACAGCAGCAACAACATGTTTCATGATGATGATGATGATGATGATGATG 375
106 roHISpPOTyrgLInGInuValSerAlaLeuGlyGlyIlePOTyriser 122
|||||
376 CTCACCCATATGACCAAGAGCTGTCTGCTTGGGTGGCAATCCCAATAC 425

123 GlnIleTyrgLInPOTyrgValAsnPheGlyValIleAspGluArle 139
|||||
426 CAATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 475
139 uHISArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleA 156
|||||
476 CCATAGAGATAGGAGATAGGAGAGAGTACTATAGGAACTCAACATAG 525
156 IAProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisGln 172
|||||
526 CTCGAGCAGAGAGATGTTATAGTTGGCAGGTTCCACACACACACCA 575
173 AlArPrArgGluGluProTyrIleHisIleAlaProGlnGlyCysGly 188
|||||
576 GCCTGGAGGAGAGAGCCCTGGATCCACCATGACACCAAGGTTGTGCA 623

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051314

seq_documentation_block:

ID AA051314 standard; DNM: 711 BP.

XX AA051314;

XX 08-DEC-1993 (first entry)

XX Encodes Asp-53 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
KW ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

XX OS

XX key Location/Qualifiers

FT misc_difference 153..155

FT /tag= a

XX /note= "wild-type GTC(Val) mutated to GAT(Asp)"

PN MO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92MO-EP03016.

XX 31-DEC-1991; 91IT-0MI3513.

XX (BIOC-) BIOCINE SCLAVO SPA.

XX Domenighini M, Hol W, Pizze M, Rappuoli R;

XX WPI; 1993-227320/28.

XX P-PSDB; AAR38730.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX - useful as vaccines against infection by Vibrio cholerae and

XX enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-Q51326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051314 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 235 A; 122 C; 156 G; 198 T; 0 other;

alignment_scores:

Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA051314 ..

Align seg 1/1 to: AA051314 from: 1 to: 711

103 ValTyrSerProHisPOTyrgLInGInuValSerAlaLeuGlyIle 119
|||||
295 GTATCAGACCCCTCACCATATGACAGAGGTTTGTGCTTAGGTGGAAT 344
119 ePOTyriserGlnIleTyrgLInPOTyrgValAsnPheGlyValIleA 136
|||||
345 ACCATATCTCTCAGATATATGATGATGATGATGATGATGATGATGATG 394
136 SPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGAACGATTACATCTGTAACAGGAAATATAGACCGGTATATACAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169
|||||
445 CTGAATATAGCTCCGCGCAGAGAGATGTTTACAGATTAGCAGGTTTCCAC 494
169 oAspHisGlnAlaTrpArgGluGluProTyrIleHisIleAlaProGln 186
|||||
495 GGATCACCAGAGTTGGAGAGAAAGAACCTGATTCATCATGACCAAG 544
186 IyCysGly 188
|||||
545 GTTGTGCA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051315

seq_documentation_block:

ID AA051315 standard; DNM: 711 BP.

XX AA051315;


```

DT 08-DEC-1993 (first entry)
DE Encodes Glu-53 E.coli heat labile toxin subunit A.
XX
XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protome A; site-directed mutagenesis; reduced toxicity;
XX ADP-ribosyltransferase activity; ss.
XX Escherichia coli.
OS
FH Key Location/Qualifiers
FT misc_difference 153..155
   /tag= a
   /note= "wild-type GTC(Val) mutated to GAA(Glu)"
PN MO9313202-A.
XX
XX 08-JUL-1993.
PD
XX 30-DEC-1992; 92WO-EP03016.
PF
XX 31-DEC-1991; 91TT-OMI3513.
PR (BIOC-) BIOCINE SCLAVO SPA.
PA
XX Domenighini M, Hol W, Pizza M, Rappuoli R;
PI MPI: 1993-227320/28.
DR P-PSDB; AAR38731.
XX
XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PR enterotoxin producing Escherichia coli
XX
XX Claim 3; Fig 2 and Page 46; 60pp; English.
PS
XX The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051315 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).
XX
SQ Sequence 711 BP; 236 A; 122 C; 156 G; 197 T; 0 other;

alignment_scores:
      Quality: 86.00      Length: 86
      Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051315 ..
Align seg 1/1 to: AA051315 from: 1 to: 711

103 ValTyrSerProHisProTyrGluGlnGluValSerAlaLeuGlyIle 119
   |||||||
295 GTATACAGCCCTCACCCATATGAACAGAGGTTCTCGTAGTGGAAT 344
   |||||||
119 eProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhcglyValIleA 136
   |||||||
345 ACCGATATTCAGATATATGATGATCGTAAATTTTGGTGTATG 394
   |||||||
136 SPGIUAAGLEuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
   |||||||
395 ATGACGATATACATCGTACACAGGAATATAGAGCCGTAATACAGAAAT 444
   |||||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169

```

```

|||||
445 CTGATATACCTCCGACAGAGATGCTTACAGATTACAGCTTCCACC 494
169 oAspHisGlnAlaTyrArgGluGlnProTyrPheHisHisAlaProGln 186
   |||||||
495 GATACACCAAGCTTGGAGAGAGAACCTCGATTATCATGACACACAG 544
186 LysCysGly 188
   |||||||
545 GTTGTGCA 552

seq_name: /STD5/gcgcdata/geneseq/geneseqn_emb1/NA1993.DAT:AA051316
seq_documentation_block:
ID AA051316 standard; DNA; 711 BP.
XX
XX AA051316;
AC
XX
XX 08-DEC-1993 (first entry)
DT
XX
XX Encodes Tyr-53 E.coli heat labile toxin subunit A.
DE
XX
XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protome A; site-directed mutagenesis; reduced toxicity;
XX ADP-ribosyltransferase activity; ss.
XX Escherichia coli.
OS
FH Key Location/Qualifiers
FT misc_difference 153..155
   /tag= a
   /note= "wild-type GTC(Val) mutated to TAC(Tyr)"
PN MO9313202-A.
XX
XX 08-JUL-1993.
PD
XX 30-DEC-1992; 92WO-EP03016.
PF
XX 31-DEC-1991; 91TT-OMI3513.
PR (BIOC-) BIOCINE SCLAVO SPA.
PA
XX Domenighini M, Hol W, Pizza M, Rappuoli R;
PI MPI: 1993-227320/28.
DR P-PSDB; AAR38732.
XX
XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PR enterotoxin producing Escherichia coli
XX
XX Claim 3; Fig 2 and Page 46; 60pp; English.
PS
XX The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051316 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).
XX
SQ Sequence 711 BP; 235 A; 123 C; 155 G; 198 T; 0 other;

alignment_scores:
      Quality: 86.00      Length: 86
      Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```


alignment_block:
US-09-528-682-1 x AAQ51316 ..

Align seg 1/1 to: AAQ51316 from: 1 to: 711

```

103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyI119
    |||
295 GTATACAGCCCTCACCATATGAAACAGAGGTTCTCGCTTAGTGGAAT 344
    |||
119 eProTyrSerGlnIleTyrGlyTyrArgValAsnPhcGlyVal111lea 136
    |||
345 ACCATATTCTCAGATATATGATGATCGTAAATTTGGTGTGATWG 394
    |||
136 SPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
    |||
395 ATGACACATTATCATCGTACACGGGAAATATAGACCCGCTATACAGAAAT 444
    |||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169
    |||
445 CTGAATATAGCTCCGGCAGAGATGTTACAGATTACAGAGTTCCACACC 494
    |||
169 oAspHisGlnAlaTrrParGluGluProTrrPLeHisHisAlaProGlnG 186
    |||
495 GGATCACCAAGCTTGAGAGAGAAACCCGTGATTCATCATGACACCAAG 544
    |||
186 LysGly 188
    |||
545 GTTGTGGA 552

```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51317

seq_documentation_block:

ID AAQ51317 standard: DNA; 711 BP.

AC AAQ51317:

DT 08-DEC-1993 (first entry)

DE Encodes Lys-63 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KM ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

OS

FT key Location/Qualifiers

FT misc_difference 184..186

FT /tag= a

FT /note= "wild-type TCT(Ser) mutated to AAG(Lys)"

PN WO9313202-A.

XX 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-OMI3513.

PA (BIOC-) BIOCINE SCLAVO SPA.

PI Domenighini M, Hol W, Pizsa M, Rappuoli R;

DR WPI: 1993-227320/28.

DR P-PSDB; AAR44016.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli

PS Claim 3: Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile

CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AAQ51314-Q51326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AAQ51317 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 236 A; 122 C; 157 G; 196 T; 0 other;

alignment_scores:

Quality:	86.00	Length:	86
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:
US-09-528-682-1 x AAQ51317 ..

Align seg 1/1 to: AAQ51317 from: 1 to: 711

```

103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyI119
    |||
295 GTATACAGCCCTCACCATATGAAACAGAGGTTCTCGCTTAGTGGAAT 344
    |||
119 eProTyrSerGlnIleTyrGlyTyrArgValAsnPhcGlyVal111lea 136
    |||
345 ACCATATTCTCAGATATATGATGATCGTAAATTTGGTGTGATWG 394
    |||
136 SPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
    |||
395 ATGACACATTATCATCGTACACGGGAAATATAGACCCGCTATACAGAAAT 444
    |||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169
    |||
445 CTGAATATAGCTCCGGCAGAGATGTTACAGATTACAGAGTTCCACACC 494
    |||
169 oAspHisGlnAlaTrrParGluGluProTrrPLeHisHisAlaProGlnG 186
    |||
495 GGATCACCAAGCTTGAGAGAGAAACCCGTGATTCATCATGACACCAAG 544
    |||
186 LysGly 188
    |||
545 GTTGTGGA 552

```

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51318

seq_documentation_block:

ID AAQ51318 standard: DNA; 711 BP.

AC AAQ51318:

DT 08-DEC-1993 (first entry)

DE Encodes Lys-97 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KM ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

OS

FT key Location/Qualifiers

FT misc_difference 277..279

FT /tag= a

FT /note= "wild-type GTT(Val) mutated to AAG(Lys)"

PN WO9313202-A.

XX 08-JUL-1993.

PF 30-DEC-1992; 92WC-EP03016.
 XX 31-DEC-1991; 91IT-0MI3513.
 XX (BIOC-) BIOCINE SCLAVO SPA.
 PA Domenighini M, Hol W, Pizza M, Rappuoli R;
 XX WPI; 1993-227320/28.
 DR P-PSDB; AAR44017.
 XX
 PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.
 XX
 CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051318 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 236 A; 123 C; 156 G; 196 T; 0 other;

alignment_scores:
 Quality: 86.00 Length: 86
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AA051318 ..

Align seg 1/1 to: AA051318 from: 1 to: 711

```

103 ValTyrSerProHisProTyrGluGlnGluValSerAlaLeuGlyIY11 119
|||||
295 GTATACAGCCCTCACCCTATATGACAGAGGTTCTCGGTAGGTGAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhenGlyValIleA 136
|||||
345 ACCATATTCCTCAGATATATGATGATGATCGTGTATTTTGGTGTGATG 394
|||||
136 SPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACAGATTACATCGTAAACAGAGGAAATAGAGCCGATATACAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169
|||||
445 CTGAATATATAGCTCCGCGAGAGATGTTACAGATTAGCAGCGTTTCCAC 494
|||||
169 oAspHisGlnAlaTrpArgGluGluProTyrIleHisIleAspProGing 186
|||||
495 GGATACACCAAGCTTGAGAGAGAAACCTGGATTATCATCAGCACCAAG 544
|||||
186 IYCYsGlyI 188
|||||
545 GTTGTGGA 552

```

seq_name: /SIDS5/9cgcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051319

seq_documentation_block:

ID AA051319 standard; DNA; 711 BP.
 XX
 AC AA051319;
 XX
 DT 08-DEC-1993 (first entry)

XX
 DE Encodes Tyr-97 E.coli heat labile toxin subunit A.
 XX
 KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
 KW protomer A; site-directed mutagenesis; reduced toxicity;
 KW ADP-ribosyltransferase activity; ss.
 XX
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT misc_difference 277..279
 FT /tag= a
 FT /note= "wild-type GTT(Val) mutated to TAC(Tyr)"
 XX
 PN WO9313202-A.
 XX
 PD 08-JUL-1993.
 XX
 PF 30-DEC-1992; 92WC-EP03016.
 XX
 PR 31-DEC-1991; 91IT-0MI3513.
 XX
 PA (BIOC-) BIOCINE SCLAVO SPA.
 XX
 PI Domenighini M, Hol W, Pizza M, Rappuoli R;
 XX WPI; 1993-227320/28.
 DR P-PSDB; AAR44018.
 XX
 PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.
 XX
 CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-051326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051319 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 235 A; 124 C; 155 G; 197 T; 0 other;

alignment_scores:
 Quality: 86.00 Length: 86
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AA051319 ..

Align seg 1/1 to: AA051319 from: 1 to: 711

```

103 ValTyrSerProHisProTyrGluGlnGluValSerAlaLeuGlyIY11 119
|||||
295 GTATACAGCCCTCACCCTATATGACAGAGGTTCTCGGTAGGTGAT 344
|||||
119 eProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhenGlyValIleA 136
|||||
345 ACCATATTCCTCAGATATATGATGATGATCGTGTATTTTGGTGTGATG 394
|||||
136 SPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACAGATTACATCGTAAACAGAGGAAATAGAGCCGATATACAGAAAT 444
|||||
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169
|||||

```



```

445 CTGATATAGCTCGGCGAGAGATGGTTACAGATTAGCAGGTTTCCACC 494
169 OASPHISGLNALATRPARGIUGLUPROTPRIEHSIALAPROGING 186
    |||||||
495 GGATCACCACCTTGAGAGAAACCTCGATTATCATGCACACCAAG 544
186 IYCYSGLY 188
    |||||||
545 GTTGTGGA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ42768

seq_documentation_block:
ID   AAQ42768 standard; DNA; 711 BP.
XX
AC   AAQ42768;
XX
DT   08-DEC-1993 (first entry)
XX
DE   E.coli heat labile toxin subunit A coding sequence.
XX
KW   enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KM   protomer A; site-directed mutagenesis; reduced toxicity;
XX   ADP-ribosyltransferase activity; ss.
XX
OS   Escherichia coli.
FH   key
FT   mat_peptide
FT   1..708
FT   /*tag= a
FT   /note= "LT-A"
XX
PN   WO9313202-A.
XX
PD   08-JUL-1993.
XX
PF   30-DEC-1992; 92WO-EP03016.
XX
PR   31-DEC-1991; 91IT-0M13513.
XX   (BIOC-) BIOGINE SCLAVO SPA.
XX   PA
XX   PI Domenighini M, Hol W, Piazza M, Rappuoli R;
XX   WPI: 1993-227320/28.
XX   DR P-PADB; AAR38728.
XX
PT   Immunogenic detoxified mutant cholera toxin and heat labile toxin
PT   - useful as vaccines against infection by Vibrio cholerae and
PT   enterotoxin producing Escherichia coli
XX
PS   Disclosure: Fig 2; 60pp; English.
XX
XX   This sequence encodes the A subunit of the heat labile toxin (LT-A)
XX   of a strain of E.coli known to affect humans. The sequence was
XX   published by Yamamoto et al. J Biol. Chem. 259, 5037-5044.
XX   Mutations at selected positions within this sequence have been found
XX   to reduce toxicity (see AAQ51314-Q51326). The invention relates to
XX   immunogenic, detoxified LT-A proteins and their use in vaccines to
XX   protect against enterotoxigenic E.coli.
XX
SQ   Sequence 711 BP; 234 A; 123 C; 156 G; 198 T; 0 other;

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAQ42768 ..
Align seg 1/1 to: AAQ42768 from: 1 to: 711

```

```

103 VALTYSERPRONISPROTYRGLUGLNUVALSERIALALEUGLYYL 119
    |||||||
295 GTATACACCCCTCACCCATATGACAGAGGTTTCTCGTTAGGTGGAAT 344
119 EPROTYSERGINILETYRGLYTRPTFRARGVALASNPHEGYVALIIEA 136
    |||||||
345 ACCATATTCTCGATATATGAGTGTATCGTTTAAATTTGGTGTGATTTG 394
136 SPGLUARGLEUNHISARGASNAARGLUYRARGSPARGTYRTRARGAN 152
    |||||||
395 ATGACGATTTACATGCTAACAGGGAATATAGACACGGTATATACGAAT 444
153 LEUASNILEALPRQALAGLUASPGLYTRARGLEUALAGLYPHERPTR 169
    |||||||
445 CTGATATAGCTCGGCGAGAGATGGTTACAGATTAGCAGGTTTCCACC 494
169 OASPHISGLNALATRPARGIUGLUPROTPRIEHSIALAPROGING 186
    |||||||
495 GGATCACCACCTTGAGAGAAACCTCGATTATCATGCACACCAAG 544
186 IYCYSGLY 188
    |||||||
545 GTTGTGGA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAV81595

seq_documentation_block:
ID   AAV81595 standard; DNA; 711 BP.
XX
AC   AAV81595;
XX
DT   11-MAY-1999 (first entry)
XX
DE   E. coli heat labile toxin gene.
XX
XX   A subunit; heat labile toxin; ADP-ribosylation; mutant; detoxification;
XX   parenteral adjuvant; antigen; immunisation; humoral response;
XX   cell-mediated immune response; virus; bacterium; parasite; fungus;
XX   tumour; allergen; pathogen; AIDS; autoimmune disease; cancer; antibody;
XX   systemic lupus erythematosus; Alzheimer's disease; diagnosis; ss.
XX
OS   Escherichia coli.
XX
PN   WO9842375-A1.
XX
PD   01-OCT-1998.
XX
PF   19-MAR-1998; 98MO-US05454.
XX
PR   18-MAR-1998; 98US-0044696.
XX   21-MAR-1997; 97US-0041227.
XX
XX   (CHIR ) CHIRON CORP.
XX
PI   Barchfeld G, Del Giudice G, Rappuoli R;
XX   WPI: 1999-070064/06.
XX   DR P-PADB; AAW67772.
XX
PT   Detoxified mutants of bacterial ADP-ribosylating toxins as
PT   parenteral adjuvants - useful to enhance humoral and cell-mediated
PT   immune responses in vertebrates when administered with selected
PT   antigen e.g. in disease treatment
XX
PS   Disclosure: Fig 1A-B; 51pp; English.
XX
XX   This sequence corresponds to the coding region for the A subunit of the
XX   E. coli heat labile toxin, an example of a bacterial ADP-ribosylating
XX   toxin. A mutant detoxified form of this protein is used in a parenteral
XX   adjuvant composition, which comprises the detoxified protein, at least
XX   one selected antigen and optionally a pharmaceutically acceptable
XX   (optionally topical) vehicle. The adjuvant composition can be

```


CC administered parenterally in conjunction with at least one antigen in
CC methods to immunise vertebrate subjects. The adjuvant has the ability
CC to enhance the humoral and cell-mediated immune responses elicited by
CC the antigen (e.g. by making the antigen more strongly immunogenic or
CC necessitating fewer/lower antigen doses). It can be administered
CC prior/subsequent to the antigen, and is preferably administered within
CC a short space of time to the same site; it can also be administered in
CC isolation from antigens as a boost following systemic or mucosal antigen
CC administration. Most preferably, the adjuvant is co-administered with
CC the antigen in the compositions and a pharmaceutically acceptable
CC carrier. The antigen may be derived from viruses, bacteria, parasites
CC and fungi or may be tumour antigens, self-antigens and allergens. The
CC compositions are therefore useful in the treatment and prevention of
CC e.g. viral diseases, allergic manifestations, diseases caused by
CC pathogens (e.g. bacteria or parasites), AIDS, autoimmune diseases
CC (e.g. Systemic Lupus Erythematosus), Alzheimer's disease and cancers.
CC The adjuvant can also be used to prepare antibodies against selected
CC antigen(s), useful e.g. for diagnostic purposes or for antigen
CC purification.

SO Sequence 711 BP; 234 A; 123 C; 156 G; 198 T; 0 other;

alignment_scores:
Quality: 86.00 Length: 86
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAV81595 ..

Align seg 1/1 to: AAV81595 from: 1 to: 711

103 ValTyrSerProHisProTyrGluGluValSerAlaLeuGlyIyl 119
|||||
295 GATACAGCCCTCAGCCATATGACAGAGAGTTCTCGATTAGTGAAT 344
119 eProTyrSerGlnIleTyrGlyTyrTyrArgValAsnPhcGlyValIleA 136
|||||
345 ACCATATTCACAGATATATGATGATGATGATGATGATGATGATGATG 394
136 SPGLuATGLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGAGACCATATACATCGTACAGGAGATATAGAGCCGTTATACAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgGluAlaGlyPheProP 169
|||||
445 CTGAATATACCTCCGCGAGAGATGTTACAGATTACAGATTCCCAAC 494
169 OASPHISGlnAlaLeuTyrArgGluGluProTyrIleHisAlaProGln 186
|||||
495 GGATCACCAAGCTTGAGAGAGAACCCGTGATTCATCATGACACCAAG 544
186 LysGlyGly 188
|||||
545 GTTGTGGA 552

seq_name: /sids5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:AA051321

seq_documentation_block:
ID AA051321 standard; DNA; 711 BP.

XX AA051321;

DT 08-DEC-1993 (first entry)

XX Encodes Lys-104 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
KW protomer A; site-directed mutagenesis; reduced toxicity;
XX ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

XX Key Location/Qualifiers
FH misc_difference 298..300
FT /tag= a
FT /note= "wild-type Trc(Tyr) mutated to AaG(Lys)"

XX W09313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92MO-EP03016.

XX 31-DEC-1991; 91IT-OM13513.

XX (BIOC-) BIOCINE SCLAVO SPA.

XX Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX WPI; 1993-227320/28.

XX P-PSDB; AAR44020.

PT - useful as vaccines against infection by Vibrio cholerae and
PT enterotoxin producing Escherichia coli

PS Claim 3; Fig 2 and Page 46; 60pp; English.

CC The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol.Chem., 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-051326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051321 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (amino
CC acid numbering is based on the cholera toxin A subunit sequence).

SO Sequence 711 BP; 235 A; 122 C; 157 G; 197 T; 0 other;

alignment_scores:
Quality: 84.00 Length: 84
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051321 ..

Align seg 1/1 to: AA051321 from: 1 to: 711

105 SerProHisProTyrGluGluGluValSerAlaLeuGlyIleProTy 121
|||||
301 AGCCCTCACCCATATGACAGAGAGTTCTCGCTTAGTGAATACATA 350
121 rSerGlnIleTyrGlyTyrTyrArgValAsnPhcGlyValIleAspGlu 138
|||||
351 TTCTCAGATATATGATGATGATGATGATGATGATGATGATGATGATG 400
138 rGluLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 154
|||||
401 GATTACATCGTACACAGGAGATATAGAGCCGTTATACAGAAATCTGAT 450
155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAsp 171
|||||
451 ATGACTCCGCGCAGAGATGTTACAGATTAGCAGATTCCACCGGATCA 500
171 sGlnAlaLeuTyrArgGluGluProTyrIleHisAlaProGlnGlyCys 188
|||||
501 CCAAGCTTGAGAGAGAACCCGTGATTCATCATGACACCAAGGTGTG 550
188 Lys 188
||

551 GA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051322

seq_documentation_block:

ID AA051322 standard; DNA; 711 BP.

AC AA051322;

DT 08-DEC-1993 (first entry)

DE Encodes Asp-104 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KM protomer A; site-directed mutagenesis; reduced toxicity;

OS Escherichia coli.

FH Key Location/Qualifiers

FT misc_difference 298..300

FT /*tag= a /note= "wild-type TAC(Tyr) mutated to GAT(Asp)"

XX MO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-OMI3513.

XX (BIOC-) BIOCINE SCLAVO SPA.

XX Domeni1ghni M, Hol W, Piazza M, Rappuoli R;

XX MPI: 1993-227320/28.

XX P-PSDB; AAR44021.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX PT - useful as vaccines against infection by Vibrio cholerae and

XX PS enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 46; 60pp; English.

XX The wild-type sequence coding for the A subunit of the heat labile

XX toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto

XX et al, J.Biol.Chem., 259, 5037-5044 - see AA042768) was subjected to

XX site-directed mutagenesis. Certain mutations were found to reduce

XX toxicity (see AA051314-051326). The invention relates to

XX immunogenic, detoxified LT-A proteins and their use in vaccines to

XX protect against enterotoxigenic E.coli. Sequence AA051322 is a

XX combination of the wild-type coding sequence and the mutagenic

XX CC primer sequence used to introduce the preferred mutation. (Amino

XX acid numbering is based on the cholera toxin A subunit sequence).

SQ Sequence 711 BP; 234 A; 122 C; 157 G; 198 T; 0 other;

alignment_scores:

Quality: 84.00

Ratio: 1.000

Percent Similarity: 100.000

Percent Identity: 100.000

Length: 84

Gaps: 0

Gaps: 0

alignment_block:

US-09-528-682-1 x AA051322 ..

Align seg 1/1 to: AA051322 from: 1 to: 711

105 SerProHisProTyrGluGluValSerAlaLeuGlyGlyProTyr 121
 |||||||
 301 AGCCCTCACCACATATGACAGAGAGGTTTCGCTTAGTGGAATACATA 350

121 rSerGlnIleTyrGlyTyrTyrArgValAsnPhnGlyValIleAspGluA 138
 |||||||
 351 TTCTCAGATATATGATGATGATATCGTGTATTGTTGGTGTATGATGATAC 400
 |||||||
 138 rGleuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
 |||||||
 401 GATTACATCGTAAACAGAGGAATATAGACCCGGTATTCAGAAATCTGAAT 450
 |||||||
 155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPheAspPhe 171
 |||||||
 451 ATAGCTCCGGCAGAGATGGTTACAGATTAGCAGGTTCCACCGGATCA 500
 |||||||
 171 sGlnAlaTyrArgGluGluProTyrPheHisAlaProGlnGlyCysG 188
 |||||||
 501 CCAAGCTTGAGAGAGAACACCGTGTATTCATCATGACCAAGGTTGTG 550
 |||||||
 188 Ly 188
 ||
 551 GA 552

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051323

seq_documentation_block:

ID AA051323 standard; DNA; 711 BP.

XX AA051323;

XX 08-DEC-1993 (first entry)

XX Encodes Ser-104 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

XX KM protomer A; site-directed mutagenesis; reduced toxicity;

XX ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

XX OS

XX FH

XX Key

XX Location/Qualifiers

XX misc_difference 298..300

XX /*tag= a

XX /note= "wild-type TAC(Tyr) mutated to AGT(Ser)"

XX MO9313202-A.

XX 08-JUL-1993.

XX 30-DEC-1992; 92WO-EP03016.

XX 31-DEC-1991; 91IT-OMI3513.

XX (BIOC-) BIOCINE SCLAVO SPA.

XX Domeni1ghni M, Hol W, Piazza M, Rappuoli R;

XX MPI: 1993-227320/28.

XX P-PSDB; AAR44022.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX PT - useful as vaccines against infection by Vibrio cholerae and

XX PS enterotoxin producing Escherichia coli

Claim 3; Fig 2 and Page 46; 60pp; English.

The wild-type sequence coding for the A subunit of the heat labile

toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto

et al, J.Biol.Chem., 259, 5037-5044 - see AA042768) was subjected to

site-directed mutagenesis. Certain mutations were found to reduce

toxicity (see AA051314-051326). The invention relates to

immunogenic, detoxified LT-A proteins and their use in vaccines to

protect against enterotoxigenic E.coli. Sequence AA051323 is a

combination of the wild-type coding sequence and the mutagenic

CC primer sequence used to introduce the preferred mutation. (Amino

acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 234 A; 122 C; 157 G; 198 T; 0 other;

alignment_scores:
Quality: 84.00 Length: 84
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051323 ..

Align seg 1/1 to: AA051323 from: 1 to: 711

```

105 SerProHisProTyrGluGlnGluValSerAlaLeuGlyGlyIleProTyr 121
|||||
301 ACCCCCTACCCATATGACAGAGAGGTTCTGCGTGAAGGAAATACCAATA 350
121 rSerGlnIleTyrGlyTyrTyrArgValAsnPhenGlyValIleAspGluA 138
|||||
351 TTCTCGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
138 rGluHisArgAsnArgGluTyrTyrArgAspArgTyrTyrArgAsnLeuAsn 154
|||||
401 GATTACATGCTAACAGGAAATATAGACCGGATATACAGAAATCTGAAT 450
155 lIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspH 171
451 ATAGCTCCGCGCAGAGGATGCTTACAGATTAGCAGGTTTCCCGCCGATCA 500
171 sGlnAlaTyrArgGluGluProTyrPheHisHisAlaProGlnGlyCysG 188
|||||
501 CCAAGCTTGAGAGAGAACCTGATCATCATCATGACCAACAAGGTTGGTG 550
188 lY 188
||
551 GA 552

```

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051324

seq_documentation_block:

ID AA051324 standard; DNA: 711 BP.

XX AA051324:

DT 08-DEC-1993 (first entry)

DE Encodes Ser-106 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

XX Key Location/Qualifiers

FT misc_difference 304..306

FT /note="wild-type CCR(Pro) mutated to AGT(Ser)"

PN W09313202-A.

PD 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-OMI3513.

PA (BIOC-) BIOGINE SCLAVO SPA.

PI Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX WPI; 1993-227320/28.

DR P-PSDB: AAR44023.

XX Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

XX Claim 3; Fig 2 and Page 46; 60pp; English.

CC The wild-type sequence coding for the A subunit of the heat labile
CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
CC site-directed mutagenesis. Certain mutations were found to reduce
CC toxicity (see AA051314-Q51326). The invention relates to
CC immunogenic, detoxified LT-A proteins and their use in vaccines to
CC protect against enterotoxigenic E.coli. Sequence AA051324 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation. (Amino
CC acid numbering is based on the cholera toxin A subunit sequence).

XX Sequence 711 BP; 235 A; 122 C; 157 G; 197 T; 0 other;

alignment_scores:
Quality: 82.00 Length: 82
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051324 ..

Align seg 1/1 to: AA051324 from: 1 to: 711

```

107 HisProTyrGluGlnGluValSerAlaLeuGlyGlyIleProTyrSerG 123
|||||
307 CACCCATATGACAGAGAGGTTCTGCGTGAAGGAAATACCAATATCTCA 356
123 nIleTyrGlyTyrTyrArgValAsnPhenGlyValIleAspGluArgLeu 140
|||||
357 GATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 406
140 lAsArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAla 156
|||||
407 ATCGTACACAGGAAATATAGACCGGATATACAGAAATCTGAATATAGCT 456
157 ProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspHisGlnAl 173
|||||
457 CCGGCGAGAGGATGTTACAGATTCAGCAGGTTTCCCGCCGATCACCACA 506
173 aTyrArgGluGluProTyrPheHisHisAlaProGlnGlyCysGly 188
|||||
507 TTGGAGAGAGAAACCTGATTCATCATCATGACCAACAAGGTTGGTGA 552

```

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051320

seq_documentation_block:

ID AA051320 standard; DNA: 711 BP.

XX AA051320:

DT 08-DEC-1993 (first entry)

DE Encodes Glu-107 E.coli heat labile toxin subunit A.

KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

OS Escherichia coli.

XX Key Location/Qualifiers

FT misc_difference 307..309

FT /note="wild-type CAC(His) mutated to GAG(Glu)"

XX MO9313202-A.
 PN 08-JUL-1993.
 XX
 PD 30-DEC-1992; 92WO-EP03016.
 XX
 PF 31-DEC-1991; 91IT-0MI3513.
 XX
 PR (BIOC-) BIOICINE SCLAVO SPA.
 XX
 PA Domenighini M, Hol W, Piazza M, Rappuoli R;
 PI WPI: 1993-227320/28.
 DR P-PSDB: AAR44019.
 DR
 XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.
 XX
 CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-Q51326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051320 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 234 A; 121 C; 158 G; 198 T; 0 other;

alignment_scores:
 Quality: 81.00 Length: 81
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AA051320 ..

Align seg 1/1 to: AA051320 from: 1 to: 711

108 PRTYRGLUGLNGLVALSERALALEUGLYLLEPRTYRSGINII 124
 ||||||||||||||||||||||||||||||||||||||||||||
 310 CCATATGACACAGAGAGGTTCTCGCTTAGGTGAAATACCATATTCACAGAT 359
 124 ERYRGLYTPYRATGVALASNPHEGLYVALILEASPGIUAIRLEUNHISA 141
 ||||||||||||||||||||||||||||||||||||||||||||
 360 AATATGATGATGATACGTGTAATTTTGTGTGATGATGAACGATTTACATC 409
 141 TGAASATRGGLUTYRATGASPARGTYRATGASNPHEUNHISALAPRO 157
 ||||||||||||||||||||||||||||||||||||||||||||
 410 GTTACACGGGAATATACAGACCGGTATTTACAGAAATCTGAATATAGCTCCG 459
 158 AIALGLUASPGIYTYRATGLENALAGLYPHEPTROASPHISGLNALATR 174
 ||||||||||||||||||||||||||||||||||||||||||||
 460 GCAGACGATGGTTACAGATTAGCAGGTTCCCGCGGATCACCAACCTTG 509
 174 PARGGLUGLUPRTGTPRIEHSIALAPROGLNGLYCYSGLY 188
 ||||||||||||||||||||||||||||||||||||||||||||
 510 GAGAGAGAAGAACCTTGATTCATCATCCACACAAAGTTGTGGA 552

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.AA051325

seq_documentation_block:

ID AA051325 standard; DNA; 711 BP.

XX AA051325;
 AC
 XX

DT 08-DEC-1993 (first entry)
 XX
 DE Encodes Glu-114 E.coli heat labile toxin subunit A.
 XX
 KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;
 KW protomer A; site-directed mutagenesis; reduced toxicity;
 KW ADP-ribosyltransferase activity; ss.
 XX
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT misc_difference 328..330
 FT /*tag=a
 FT /note="wild-type Tcr(Ser) mutated to GAA(Glu)"
 XX
 PN WO9313202-A.
 XX
 PD 08-JUL-1993.
 XX
 PF 30-DEC-1992; 92WO-EP03016.
 XX
 PR 31-DEC-1991; 91IT-0MI3513.
 XX
 PA (BIOC-) BIOICINE SCLAVO SPA.
 XX
 PI Domenighini M, Hol W, Piazza M, Rappuoli R;
 XX
 DR WPI: 1993-227320/28.
 DR P-PSDB: AAR44024.
 DR
 XX Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.
 XX
 CC The wild-type sequence coding for the A subunit of the heat labile
 CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto
 CC et al, J.Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to
 CC site-directed mutagenesis. Certain mutations were found to reduce
 CC toxicity (see AA051314-Q51326). The invention relates to
 CC immunogenic, detoxified LT-A proteins and their use in vaccines to
 CC protect against enterotoxigenic E.coli. Sequence AA051325 is a
 CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation. (Amino
 CC acid numbering is based on the cholera toxin A subunit sequence).
 XX
 SO Sequence 711 BP; 236 A; 122 C; 157 G; 196 T; 0 other;

alignment_scores:
 Quality: 74.00 Length: 74
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AA051325 ..

Align seg 1/1 to: AA051325 from: 1 to: 711

115 AIALGLNGLYLLEPRTYRSGINIIETRYGLYTPYRATGVALAS 131
 ||||||||||||||||||||||||||||||||||||||||||||
 331 GCGTTAGGTGGAATACCATATTTCTCAGATATATGAGATGATCGTTAA 360
 131 NPHEGLYVALILEASPGIUAIRLEUNHISATGASPARGTYRATGASPA 148
 ||||||||||||||||||||||||||||||||||||||||||||
 381 TTTTGGTGTGATGATGAGACGATTACATCTGTAACAGGAAATATAGAGACC 430
 148 TGTYTYRATGASNPHEUNHISALAPROALAGLUASPGIYTYRATGLEN 164
 ||||||||||||||||||||||||||||||||||||||||||||
 431 GGTATTTACAGAAATCTGAATATAGCTCCGCGAGAGATGCTTACAGATTA 480
 165 AIALGLYPHEPTROASPHISGLNALATRPARGLUGLUPRTGTPRIEHSI 181

|||||
 481 GCAGGTTTCCACCGGATCACCAGCTGGAGAGAAACCTGATTC A 530
 181 SHSAlaProGInGlyCysGly 188
 |||||
 531 TCATGCACCAACAAGTTGTGGA 552

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51326

seq_documentation_block:
 ID AAQ51326 standard; DNA; 711 BP.

XX
 AC AAQ51326;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-114 E.coli heat labile toxin subunit A.

XX enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX Escherichia coli.

OS
 FH Key Location/Qualifiers
 FT misc_difference 328..330

FT /tag= a /note= "wild-type TCR(Ser) mutated to AAA(Lys)"

XX
 PM W09313202-A.

XX
 PD 08-JUL-1993.

XX
 PF 30-DEC-1992; 92WO-EP03016.

XX
 PR 31-DEC-1991; 91IT-OM13513.

XX
 PA (BIOC-) BIOGINE SCLAVO SPA.

XX
 PI Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX
 DR WPI: 1993-227320/28.

XX
 DR P-PSDB; AAR44025.

XX
 PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

XX
 PS Claim 3; Fig 2 and Page 46; 60pp; English.

XX
 CC The wild-type sequence coding for the A subunit of the heat labile

CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto

CC et al, J Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to

CC site-directed mutagenesis. Certain mutations were found to reduce

CC toxicity (see AAQ51314-051326). The invention relates to

CC immunogenic, detoxified LT-A proteins and their use in vaccines to

CC protect against enterotoxigenic E.coli. Sequence AAQ51326 is a

CC combination of the wild-type coding sequence and the mutagenic

CC primer sequence used to introduce the preferred mutation. (Amino

CC acid numbering is based on the cholera toxin A subunit sequence).

XX
 SQ Sequence 711 BP; 237 A; 122 C; 156 G; 196 T; 0 other;

alignment_scores:

Quality: 74.00 Length: 74
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AAQ51326 ..
 Align seg 1/1 to: AAQ51326 from: 1 to: 711

115 AlaLeuGlyGlyLeuProTyrSerGlnIleTyrGlyTrpTyrArgValas 131
 |||||
 331 GCCTTAGGTGGATACCATATTTCTCAGATATATGATGATGTGTGA 380
 131 npheGlyValIleAspGluArgLeuHisArgAsnArgGlyTyrArgAspA 148
 |||||
 381 TTTTGGTGTGATTTGATGACGATTTACATCGTACACGAGATATATAGAGACC 430
 148 rgtTyrTyrArgAsnLeuAsnIleAlaProAlaGluAspGlyTyrArgLeu 164
 |||||
 431 GGTATTACAGAAATCTGAATATAGCTCCGGCAGAGATGGTTACAGATT A 480
 165 AlaGlyPheProAspHisGlnAlaTPrArgGluGluProTPrIleu 181
 |||||
 481 GCAGGTTTCCACCGGATCACCAGCTTGAGAGAAACCTGATTC A 530
 181 SHSAlaProGInGlyCysGly 188
 |||||
 531 TCATGCACCAACAAGTTGTGGA 552

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1995.DAT:AAQ3448

seq_documentation_block:
 ID AAT03448 standard; DNA; 777 BP.

XX
 AC AAT03448;

DT 02-MAY-1996 (first entry)

XX
 DE Shigella flexneri thermolabile enterotoxin gene fragment LTF.

XX
 KW Thermolabile enterotoxin; subunit-A; subunit-B; detection; probe;

XX
 KW LTF fragment; enterobacteria; ss.

XX
 OS Shigella flexneri.

XX
 PN RU2031948-C1.

XX
 PD 27-MAR-1995.

XX
 PF 11-DEC-1991; 91RU-0016860.

XX
 PR 11-DEC-1991; 91SU-4016860.

XX
 PA (NIZH-) NIZHEGOROD EPIDEMIOLOGY MICROBIOLOG INST.

XX
 PI Mazepa VN, Skoblo LE, Ulanova TI;

XX
 DR WPI: 1995-343066/44.

XX
 PT LTF DNA fragment for recognition of thermo:labile enterobacterium

PT enterotoxin gene - can be used as a probe for recognition of

XX
 PT LT-toxin DNA gene

XX
 PS Claim 1; Columns 3-4; 5pp; Russian.

XX
 CC The present sequence is that of the LTF fragment which encodes part

CC of the thermolabile enterotoxin subunit-A sequence and all of the

CC subunit-B sequence of Shigella flexneri. The DNA fragment is useful

CC as a probe for specific detection of toxigenic enterobacteria.

XX
 SQ Sequence 777 BP; 288 A; 141 C; 153 G; 195 T; 0 other;

alignment_scores:

Quality: 30.00 Length: 30
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AAT03448 ..

Align seg 1/1 to: AAT03448 from: 1 to: 777

173 AlarpargglugluProtrpIleHisAlaprogInglycylAs 189
 ||||||||||||||||||||||||||||||||||||||||
 3 GCTTGGAGAGAGAACCTTGATTCATCATGCACACAGAGGTGGAA 52
 189 nSerSerArgThrIleThrGlyAspThrCysAsnGluGlu 202
 ||||||||||||||||||||||||||||||||||||||||
 53 TTCATCAAGAACATTCAGGTGATCTGTGAATGAGAGAG 92

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: AAN30099

seq_documentation_block:
 ID AAN30099 standard; DNA; 795 BP.

AC AAN30099;

DT 02-JUN-1992 (first entry)

DE Sequence of HindII fragment from pEBK620 contg. an enterotoxoid gene.

XX Enterotoxoid; immunogen; vaccine; ds.

XX Key Location/Qualifiers

FT misc_feature 1..117

FT misc_feature /tag= a

FT misc_feature /label= gal e

FT misc_feature 118..213

FT misc_feature /tag= b

FT misc_feature /label= 196 bp

FT misc_feature 214..426

FT misc_feature /tag= c

FT misc_feature /label= 111 bp

FT misc_feature 427..630

FT misc_feature /tag= d

FT misc_feature /label= 196 bp

FT misc_feature 631

FT misc_feature /tag= e

FT misc_feature /label= Ent - LT

FT misc_feature 631

FT misc_feature /tag= e

FT misc_feature /label= Ent - LT

FT misc_feature 631

FT misc_feature /tag= e

FT misc_feature /label= Ent - LT

FT misc_feature 631

FT misc_feature /tag= e

FT misc_feature /label= Ent - LT

FT misc_feature 631

alignment_scores:

Quality: 30.00

Ratio: 1.000

Percent Similarity: 100.000

Percent Identity: 100.000

alignement_block:

US-09-528-682-1 x AAN30099

Align seg 1/1 to: AAN30099 from: 1 to: 795

198 ThrCysAsnGluGluThrGlnAsnLeuSerThrIleTyrLeuArgGluTyr 214
 ||||||||||||||||||||||||||||||||||||||||
 694 ACTGTGATGAGAGAGACCCAGATCTGAGCACAATATATCTCAGAGGATA 743
 214 rGlnSerLysValLysArgGlnIlePheSerAspTyrGln 227
 ||||||||||||||||||||||||||||||||||||||||
 744 TCAATCAAAAGCTTAAGCAGCATATTTTCAGACTATCAG 783

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: AAF74431

seq_documentation_block:
 ID AAF74431 standard; DNA; 347 BP.

AC AAF74431;

DT 08-MAY-2001 (first entry)

DE Vibrio cholerae ctxA gene 347 bp region nucleotide sequence.

XX Vibrio cholerae: ctxA gene; detection; identification; toxigenic; ds.

XX Vibrio cholerae.

XX ZA200001349-A.

XX 27-DEC-2000.

XX 16-MAR-2000; 2000ZA-0001349.

XX 16-MAR-2000; 2000ZA-0001349.

XX (UYPR-) UNIV PRETORIA.

XX Theron J, Venter SN, Brozel VS, Du Preez M;

XX WPI: 2001-191829/19.

XX Novel oligonucleotide primer useful for PCR amplification of toxigenic

XX Vibrio cholerae nucleic acid for detecting toxigenic Vibrio cholerae

XX organisms present in the sample -

XX Example 2; Fig 7; 44p; English.

XX The present invention describes a method for amplifying (a) toxigenic

XX Vibrio cholerae nucleic acid by: (a) performing a polymerase chain

XX reaction (PCR), by combining a medium containing toxigenic V. cholerae

XX nucleic acid with an oligonucleotide primer (I) which will bind to a

XX section of V. cholerae nucleic acid, in an environment suitable for

XX carrying out PCR; (b) combining the medium with a second primer which

XX will bind downstream to the first primer to a section of the

XX complementary strand of the V. cholerae nucleic acid; and (c) amplifying

XX corresponding sections of V. cholerae nucleic acid present in the medium.

XX The method is useful for detecting toxigenic V. cholerae organisms or

XX toxigenic V. cholerae nucleic acid, e.g., ctxA gene, in an environmental

XX water sample e.g., tap water sample, river water sample, lake water

XX sample or a sewage sample. The method is useful in PCR assays for

XX detecting a microorganism or nucleic acid of the microorganism in a

XX sample. (I) and amplification products are useful for detecting toxigenic

XX V. cholerae nucleic acid in a sample. The PCR method involving (I) is

XX rapid, economical, specific and sensitive for detecting toxigenic

XX V. cholerae in water samples. The present sequence represents a

XX region from the V. cholerae ctxA gene, which is used in an example from

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC the present invention.
 XX
 SQ Sequence 347 BP; 101 A; 66 C; 77 G; 103 T; 0 other;

alignment_scores:
 Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AAF74431 ..

Align seg 1/1 to: AAF74431 from: 1 to: 347

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
 ||||||||||||||||||||||||||||||||||||||||||||||||
 67 GAGTACTTTGACCGAGGACTCAATGATATCAACCTTTATGATCATGC 116
 45 aargglythrGlnThrGlyPheValArg 54
 ||||||||||||||||||||||||||||||||||||||||||||||||
 117 AAGAGGAACTCAGACGGGATTGTAGG 144

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAF74430

seq_documentation_block:

ID AAF74430 standard; DNA: 564 BP.

XX AAF74430;

DT 08-MAY-2001 (first entry)

DE Vibrio cholerae ctxA gene 564 bp region nucleotide sequence.

KW Vibrio cholerae; ctxA gene; detection; identification; toxigenic; ds.

XX OS Vibrio cholerae.

XX ZA200001349-A.

PD 27-DEC-2000.

PE 16-MAR-2000; 2000ZA-0001349.

PR 16-MAR-2000; 2000ZA-0001349.

PA (UYPR-) UNIV PRETORIA.

PI Theron J, Venter SN, Brozel VS, Du Preez M;

DR WPI; 2001-191829/19.

PT Novel oligonucleotide primer useful for PCR amplification of toxigenic
 Vibrio cholerae nucleic acid for detecting toxigenic Vibrio cholerae
 PT organisms present in the sample

PS Example 1; Fig 6; 44pp: English.

XX The present invention describes a method for amplifying (A) toxigenic
 CC Vibrio cholerae nucleic acid by: (a) performing a polymerase chain
 CC reaction (PCR), by combining a medium containing toxigenic V. cholerae
 CC nucleic acid with an oligonucleotide primer (I) which will bind to a
 CC section of V. cholerae nucleic acid, in an environment suitable for
 CC carrying out PCR; (b) combining the medium with a second primer which
 CC will bind downstream to the first primer to a section of the
 CC complementary strand of the V. cholerae nucleic acid; and (c) amplifying
 CC corresponding sections of V. cholerae nucleic acid present in the medium.
 CC The method is useful for detecting toxigenic V. cholerae organisms or
 CC toxigenic V. cholerae nucleic acid, e.g., ctxA gene, in an environmental
 CC water sample e.g., tap water sample, river water sample, lake water
 CC sample or a sewage sample. The method is useful in PCR assays for
 CC detecting a microorganism or nucleic acid of the microorganism in a
 CC sample. (I) and amplification products are useful for detecting toxigenic

CC V. cholerae nucleic acid in a sample. The PCR method involving (I) is
 CC rapid, economical, specific and sensitive for detecting toxigenic
 CC V. cholerae in water samples. The present sequence represents a
 CC region from the V. cholerae ctxA gene, which is used in an example from
 CC the present invention.

SQ Sequence 564 BP; 159 A; 105 C; 139 G; 161 T; 0 other;

alignment_scores:
 Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AAF74430 ..

Align seg 1/1 to: AAF74430 from: 1 to: 564

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
 ||||||||||||||||||||||||||||||||||||||||||||||||
 67 GAGTACTTTGACCGAGGACTCAATGATATCAACCTTTATGATCATGC 116
 45 aargglythrGlnThrGlyPheValArg 54
 ||||||||||||||||||||||||||||||||||||||||||||||||
 117 AAGAGGAACTCAGACGGGATTGTAGG 144

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ51328

seq_documentation_block:

ID AAQ51328 standard; DNA: 723 BP.

XX AAQ51328;

DT 08-DEC-1993 (first entry)

DE Encodes Lys-63 cholera toxin subunit A.

KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;
 KW protomer A; site-directed mutagenesis; reduced toxicity;

KW ADP-ribosyltransferase activity; ss.

XX OS Vibrio cholerae.

XX FH Key Location/Qualifiers

FT misc_difference 187..189

FT /*tag= a /note= "wild-type TCA(Ser) mutated to AAG(Lys)"

PN WO9313202-A.

PD 08-JUL-1993.

PE 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-0MI3513.

PA (BIOC-) BIOCINE SCLAVO SPA.

PI Domenighini M, Hol W, Pizza M, Rappuoli R;

DR WPI; 1993-227320/28.

PT P-PADB; AAR44027.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
 PT - useful as vaccines against infection by Vibrio cholerae and
 PT enterotoxin producing Escherichia coli
 XX Claim 3; Fig 2 and Page 48; 60pp: English.

XX The wild-type sequence coding for the A subunit of the cholera
 CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AAQ46318)
 CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051328 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SQ Sequence 723 BP; 228 A; 124 C; 164 G; 207 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051328 ..

Align seg 1/1 to: AA051328 from: 1 to: 723

29 GUTYTPHESPARGLYTHRGINMETASNILEASNLEUTYRASPHTSAI 45
|||||
85 GAGTACTTTGACCGAGTACTCAATGAAATATCAACCTTATGATCATGTC 134
45 AARGGLYTHRGINTHRGLYPHEVALARG 54
|||||
135 AAGAGCACTACAGACGGGATTTGTTAGG 162

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051329

seq_documentation_block:
ID AA051329 standard; DNA; 723 BP.

XX
AC AA051329;

DT 08-DEC-1993 (first entry)

XX
DE Encodes Lys-97 cholera toxin subunit A.

XX
KM Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KM ADP-ribosyltransferase activity; ss.

XX
OS Vibrio cholerae.

XX
FH Key Location/Qualifiers

FT misc_difference 289..291

FT /*tag= a /note= "wild-type GGT(Val) mutated to AAC(Lys)"

XX
PN WO9313202-A.

XX
PD 08-JUL-1993.

XX
PE 30-DEC-1992; 92WO-EP03016.

XX
PR 31-DEC-1991; 91IT-0M13513.

XX
PA (BIOC-) BIOCINE SCLAVO SPA.

XX
PI Domenighini M, Hol W, Piazza M, Rappuoli R;

XX
DR WPI: 1993-227320/28.

XX
DR P-PSDB: AAR44028.

XX
PS Claim 3; Fig 2 and Page 48; 60pp; English.

XX
CC The wild-type sequence coding for the A subunit of the cholera
CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051329 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SQ Sequence 723 BP; 229 A; 125 C; 163 G; 206 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051329 ..

Align seg 1/1 to: AA051329 from: 1 to: 723

29 GUTYTPHESPARGLYTHRGINMETASNILEASNLEUTYRASPHTSAI 45
|||||
85 GAGTACTTTGACCGAGTACTCAATGAAATATCAACCTTATGATCATGTC 134
45 AARGGLYTHRGINTHRGLYPHEVALARG 54
|||||
135 AAGAGCACTACAGACGGGATTTGTTAGG 162

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051330

seq_documentation_block:
ID AA051330 standard; DNA; 723 BP.

XX
AC AA051330;

DT 08-DEC-1993 (first entry)

XX
DE Encodes Ser-106 cholera toxin subunit A.

XX
KM Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KM ADP-ribosyltransferase activity; ss.

XX
OS Vibrio cholerae.

XX
FH Key Location/Qualifiers

FT misc_difference 316..318

FT /*tag= a /note= "wild-type CCR(Pro) mutated to AGC(Ser)"

XX
PN WO9313202-A.

XX
PD 08-JUL-1993.

XX
PE 30-DEC-1992; 92WO-EP03016.

XX
PR 31-DEC-1991; 91IT-0M13513.

XX
PA (BIOC-) BIOCINE SCLAVO SPA.

XX
PI Domenighini M, Hol W, Piazza M, Rappuoli R;

XX
DR WPI: 1993-227320/28.

XX
DR P-PSDB: AAR44029.

XX
PS Claim 3; Fig 2 and Page 48; 60pp; English.

XX
CC The wild-type sequence coding for the A subunit of the cholera
CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051330 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SQ Sequence 723 BP; 228 A; 124 C; 164 G; 207 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051330 ..

Align seq 1/1 to: AA051330 from: 1 to: 723

```
29 GluTyrPheAspArgGlyThrGlnMetAsn1LeaSnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTTATGATCATGC 134
|||||
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACTCAGACGCGGATTGTTAGG 162
```

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051331

seq_documentation_block:
ID AA051331 standard; DNA; 723 BP.

XX AC AA051331;
XX DT 08-DEC-1993 (first entry)

XX DE Encodes Asn-107 cholera toxin subunit A.

XX KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;
XX KW protomer A; site-directed mutagenesis; reduced toxicity;
XX KW ADP-ribosyltransferase activity; ss.

XX OS Vibrio cholerae.

XX FH Key Location/Qualifiers
XX FT misc_difference 319..321 /tag= a
XX FT /note= "wild-type CAT(His) mutated to AAC(Asn)"

XX PN WO9313202-A.

XX PD 08-JUL-1993.

XX PF 30-DEC-1992; 92WO-EP03016.

XX PR 31-DEC-1991; 91IT-OMI3513.

XX PA (BIOC-) BIOCINE SCLAVO SPA.

XX PI Domenighini M, Hol W, Pizsa M, Rappuoli R.

XX DR WPI; 1993-227320/28.

XX DR P-PSDB; AAR44030.

XX PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
XX PT - useful as vaccines against infection by Vibrio cholerae and
XX PT enterotoxin producing Escherichia coli

XX PS Claim 3; Fig 2 and Page 48; 60pp; English.

CC The wild-type sequence coding for the A subunit of the cholera
CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051331 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.
XX
SQ Sequence 723 BP; 228 A; 125 C; 163 G; 207 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA051331 ..

Align seq 1/1 to: AA051331 from: 1 to: 723

```
29 GluTyrPheAspArgGlyThrGlnMetAsn1LeaSnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTTATGATCATGC 134
|||||
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACTCAGACGCGGATTGTTAGG 162
```

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051332

seq_documentation_block:
ID AA051332 standard; DNA; 723 BP.

XX AC AA051332;
XX DT 08-DEC-1993 (first entry)

XX DE Encodes Ser-110 cholera toxin subunit A.

XX KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;
XX KW protomer A; site-directed mutagenesis; reduced toxicity;
XX KW ADP-ribosyltransferase activity; ss.

XX OS Vibrio cholerae.

XX FH Key Location/Qualifiers
XX FT misc_difference 328..330 /tag= a
XX FT /note= "wild-type GAA(Glu) mutated to TGC(Ser)"

XX PN WO9313202-A.

XX PD 08-JUL-1993.

XX PF 30-DEC-1992; 92WO-EP03016.

XX PR 31-DEC-1991; 91IT-OMI3513.

XX PA (BIOC-) BIOCINE SCLAVO SPA.

XX PI Domenighini M, Hol W, Pizsa M, Rappuoli R.

XX DR WPI; 1993-227320/28.

XX DR P-PSDB; AAR44031.

XX PT Immunogenic detoxified mutant cholera toxin and heat labile toxin
XX PT - useful as vaccines against infection by Vibrio cholerae and
XX PT enterotoxin producing Escherichia coli

XX PS Claim 3; Fig 2 and Page 48; 60pp; English.

CC The wild-type sequence coding for the A subunit of the cholera
CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)
CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-Q51334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051332 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.

XX
SQ Sequence 723 BP; 225 A; 126 C; 163 G; 209 T; 0 other;

alignment_scores:

Quality:	26.00	Length:	26
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x AA051332 ..

Align seg 1/1 to: AA051332 from: 1 to: 723

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACGAGGACTCAAAATGAAATATCAACCTTATGATCATGC 134
|||||
45 AARGGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACCTCAGACGGGATTGTAGG 162
```

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051333

seq_documentation_block:

ID AA051333 standard; DNA; 723 BP.

AC AA051333;

DT 08-DEC-1993 (first entry)

DE Encodes Ala-112 cholera toxin subunit A.

KM Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KM ADP-ribosyltransferase activity; ss.

OS Vibrio cholerae.

FT Key Location/Qualifiers

FT misc_difference 334..336 /tag= a

PN WO9313202-A.

PD 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-OMI3513.

PA (BIOC-) BIOCINE SCLAVO SPA.

PI Domenighini M, Hol W, Piazza M, Rappuoli R;

DR WPI: 1993-227320/28.

DR P-PSDB; AAR44032.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

PS Claim 3; Fig 2 and Page 48; 60pp; English.

CC The wild-type sequence coding for the A subunit of the cholera

CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)

CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-Q51334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051333 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.

XX
SQ Sequence 723 BP; 225 A; 126 C; 163 G; 209 T; 0 other;

alignment_scores:

Quality:	26.00	Length:	26
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x AA051333 ..

Align seg 1/1 to: AA051333 from: 1 to: 723

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACGAGGACTCAAAATGAAATATCAACCTTATGATCATGC 134
|||||
45 AARGGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACCTCAGACGGGATTGTAGG 162
```

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051334

seq_documentation_block:

ID AA051334 standard; DNA; 723 BP.

AC AA051334;

DT 08-DEC-1993 (first entry)

DE Encodes Glu-114 cholera toxin subunit A.

KM Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

KW protomer A; site-directed mutagenesis; reduced toxicity;

KM ADP-ribosyltransferase activity; ss.

OS Vibrio cholerae.

FT Key Location/Qualifiers

FT misc_difference 340..342 /tag= a

PN WO9313202-A.

PD 08-JUL-1993.

PF 30-DEC-1992; 92WO-EP03016.

PR 31-DEC-1991; 91IT-OMI3513.

PA (BIOC-) BIOCINE SCLAVO SPA.

PI Domenighini M, Hol W, Piazza M, Rappuoli R;

DR WPI: 1993-227320/28.

DR P-PSDB; AAR44032.

PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

PT - useful as vaccines against infection by Vibrio cholerae and

PT enterotoxin producing Escherichia coli

PS Claim 3; Fig 2 and Page 48; 60pp; English.

CC The wild-type sequence coding for the A subunit of the cholera

CC toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318)

CC was subjected to site-directed mutagenesis. Certain mutations were

CC found to reduce toxicity (see AA051327-051334). The invention relates
CC to immunogenic, detoxified CT-A proteins and their use in vaccines
CC to protect against Vibrio cholerae. Sequence AA051334 is a
CC combination of the wild-type coding sequence and the mutagenic
CC primer sequence used to introduce the preferred mutation.

XX
SQ Sequence 723 BP; 229 A; 124 C; 164 G; 206 T; 0 other;

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA051334 ..

Align seg 1/1 to: AA051334 from: 1 to: 723

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGAATATCATCACTTTATGATCATGTC 134

45 aArgGlyThrGlnThrGlyPheValArg 54
|||||

135 AAGAGGAACCTCAGACGCGATTGTGTAGG 162

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AA046318

seq_documentation_block:

ID AA046318 standard; DNA; 723 BP.

XX AC AA046318;

XX DT 08-DEC-1993 (first entry)

XX DE Cholera toxin subunit A coding sequence.

XX KW Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

XX KM protomer A; site-directed mutagenesis; reduced toxicity;

XX ADP-ribosyltransferase activity; ss.

XX OS Vibrio cholerae.

XX FH Key Location/Qualifiers

FT mat_peptide 1..720
FT /*tag= a
FT /note= "CT-A"

XX W09313202-A.

XX PD 08-JUL-1993.

XX PF 30-DEC-1992; 92WO-EP03016.

XX PR 31-DEC-1991; 91IT-0M13513.

XX PA (BIOC-) BIOCINE SCLAVO SPA.

XX PI Domenighini M, Hol W, Pizsa M, Rappuoli R;

XX DR WPI: 1993-227320/28.

XX P-PSDB; AAR38729.

XX PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX PT - useful as vaccines against infection by Vibrio cholerae and

XX PS enterotoxin producing Escherichia coli

XX PS Disclosure: Fig 2; 60pp; English.

CC This sequence encodes the A subunit of the cholera toxin (CT-A)
CC of Vibrio cholerae. The sequence was published by Mekalanos et al.,
CC Nature 306, 551 (1983). Mutations at selected positions within this

CC sequence have been found to reduce toxicity (see AA051327-051334). The
CC invention relates to immunogenic, detoxified CT-A proteins and their
CC use in vaccines to protect against cholera.

XX
SQ Sequence 723 BP; 227 A; 125 C; 163 G; 208 T; 0 other;

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA046318 ..

Align seg 1/1 to: AA046318 from: 1 to: 723

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGAATATCATCACTTTATGATCATGTC 134

45 aArgGlyThrGlnThrGlyPheValArg 54
|||||

135 AAGAGGAACCTCAGACGCGATTGTGTAGG 162

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV81596

seq_documentation_block:

ID AAV81596 standard; DNA; 723 BP.

XX AC AAV81596;

XX DT 11-MAY-1999 (first entry)

XX DE Cholera toxin gene.

XX KW A subunit; heat labile toxin; ADP-ribosylation; mutant; detoxification;

XX KM parenteral adjuvant; antigen; antigen; immunisation; humoral response;

XX KM cell-mediated immune response; virus; bacterium; parasite; fungus;

XX KM tumour; allergen; pathogen; AIDS; autoimmune disease; cancer; antibody;

XX KM systemic lupus erythematosus; Alzheimer's disease; diagnosis; ss.

XX OS Vibrio cholerae.

XX PN W09842375-A1.

XX PD 01-OCT-1998.

XX PF 19-MAR-1998; 98WO-US05454.

XX PR 18-MAR-1998; 98US-0044696.

XX PR 21-MAR-1997; 97US-0041227.

XX PA (CHIR) CHIRON CORP.

XX PI Barchfeld G, Del Giudice G, Rappuoli R;

XX DR WPI: 1999-070064/06.

XX PT Detoxified mutants of bacterial ADP-ribosylating toxins as

XX PT parenteral adjuvants - useful to enhance humoral and cell-mediated

XX PT immune responses in vertebrates when administered with selected

XX PS antigen e.g. in disease treatment

XX PS Disclosure: Fig 1A-B; 51pp; English.

CC This sequence corresponds to the coding region for the A subunit of the
CC cholera toxin, an example of a bacterial ADP-ribosylating toxin. A
CC mutant detoxified form of this protein is used in a parenteral adjuvant
CC composition, which comprises the detoxified protein, at least one
CC selected antigen and optionally a pharmaceutically acceptable (optionally
CC topical) vehicle. The adjuvant composition can be administered
CC parenterally in conjunction with at least one antigen in methods to

immunise vertebrate subjects. The adjuvant has the ability to enhance the humoral and cell-mediated immune responses elicited by the antigen (e.g. by making the antigen more strongly immunogenic or necessitating fewer/lower antigen doses). It can be administered prior/subsequent to the antigen, and is preferably administered within a short space of time to the same site; it can also be administered in isolation from antigens as a boost following systemic or mucosal antigen administration. Most preferably, the adjuvant is co-administered with the antigen in the compositions and a pharmaceutically acceptable carrier. The antigen may be derived from viruses, bacteria, parasites and fungi or may be tumour antigens, self-antigens and allergens. The compositions are therefore useful in the treatment and prevention of e.g. viral diseases, allergic manifestations, diseases caused by pathogens (e.g. bacteria or parasites), AIDS, autoimmune diseases (e.g. systemic lupus erythematosus), Alzheimer's disease and cancers. The adjuvant can also be used to prepare antibodies against selected antigen(s), useful e.g. for diagnostic purposes or for antigen purification.

Sequence 723 BP; 227 A; 125 C; 163 G; 208 T; 0 other;

alignment_scores: Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAV81596 ..

Align seg 1/1 to: AAV81596 from: 1 to: 723

29 GluTyRPhesPARGlyThrGlnMetAsnIleAsnLeuTyRAspHisAl 45
|||||
85 GAGTACTTTGACGAGGTACTCAATGAAATATCAACCTTATGATCATGC 134

45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
135 AAGGGAACCTCAGCGGCGATTGTGACG 162

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51157

seq_documentation_block:
ID AAA51157 standard; cDNA; 777 BP.

AC AAA51157;

DT 26-SEP-2000 (first entry)

DE Plant-optimized V. cholerae toxin A subunit coding sequence.

KW Heat-labile toxin; CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
adjuvant; anti-bacterial; ss.

XX Vibrio cholerae.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..777 /*tag= a
 /product= cholera_toxin_subunit_A

PN W0200037609-A2.

PD 29-JUN-2000.

XX 22-DEC-1999; 99WO-US30747.

XX 22-DEC-1998; 98US-0113507.

PA (BOYC-) BOYCE THOMPSON INST PLANT RES.

PA (MASO/) MASON H S.

PA (ARNT/) ARNTZEN C J.

PI Mason HS, Arntzen CJ;

XX WPI; 2000-442653/38.

DR P-PDB; AAY96653.

XX New polynucleotides encoding LT-A or CT-A polypeptides for the

PT transformation of plant cells, useful in immunogenic compositions to

PT elicit immune responses in animals

XX Disclosure; Fig 4A; 103pp; English.

XX This plant-codon optimized cDNA encodes a synthetic Vibrio cholerae
CC cholera toxin (CT) A subunit (CT-A). The sequence has no cryptic signal
CC sequence and no CG (potential methylation sites) sequences. Novel
CC polynucleotides encode a mutant Escherichia coli heat-labile toxin (LT)
CC A subunit (LT-A) polypeptide or a mutant V. cholerae cholera toxin (CT)
CC A subunit (CT-A) polypeptide, which have reduced enzyme activity as
CC compared to the wild-type LT-A or CT-A polypeptide and where at least one
CC of the codons is altered to a plant preferred codon. The polynucleotide
CC further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or
CC a CT B subunit (CT-B). The polynucleotides are useful for the
CC transformation of plant cells for the production of transgenic plants
CC to produce edible vaccines, especially oral vaccines in transgenic plants
CC for the prophylactic or therapeutic treatment against E. coli or
CC V. cholerae. The mutant polypeptides are also useful as adjuvants.

XX Sequence 777 BP; 209 A; 174 C; 191 G; 203 T; 0 other;

alignment_scores: Quality: 26.00 Length: 26
 Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAA51157 ..

Align seg 1/1 to: AAA51157 from: 1 to: 777

29 GluTyRPhesPARGlyThrGlnMetAsnIleAsnLeuTyRAspHisAl 45
|||||
139 GAGTACTTTGACAGGGGACTCAGATGAACTCAACCTTATGACACATGC 188

45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACCTCAACTGCGATTGTGACG 216

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51158

seq_documentation_block:
ID AAA51158 standard; cDNA; 777 BP.

AC AAA51158;

DT 26-SEP-2000 (first entry)

DE Plant-optimized mutant V. cholerae toxin A subunit K63 coding sequence.

KW Heat-labile toxin; CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
adjuvant; anti-bacterial; S63K; ss.

XX Vibrio cholerae.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..777 /*tag= a
 /product= mutant_cholera_toxin_subunit_A

PN W0200037609-A2.

PD 29-JUN-2000.

PF 22-DEC-1999; 99WO-US30747.
XX
PR 22-DEC-1998; 98US-0113507.
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX
PI Mason HS, Arntzen CJ;
XX
DR WPI; 2000-442653/38.
XX
XX
PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX
PS Claim 15; Fig 4B; 103pp; English.
XX
XX This plant-codon optimized cDNA encodes a mutant synthetic
CC Vibrio cholerae cholera toxin (CT) A subunit (CT-A) having a S63K
CC substitution. The sequence has no cryptic signal sequence and no CG
CC (potential methylation sites) sequences. Novel polynucleotides encode a
CC mutant Escherichia coli heat-labile toxin (LT) subunit A (LT-A)
CC polypeptide or a mutant V. cholerae cholera toxin (CT) A subunit (CT-A)
CC polypeptide, which have reduced enzyme activity as compared to the
CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
CC is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible
CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The
CC mutant polypeptides are also useful as adjuvants.
XX
SQ Sequence 777 BP; 211 A; 172 C; 192 G; 202 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAA51158 ..

Align seg 1/1 to: AAA51158 from: 1 to: 777

29 GluYrPheAspArgGlyThrGlnMetAsnIleAsnLeuYrAspHisAl 45
|||||
139 GAGTACTTGTGACAGGGGTACTGACATGCAACCTTTATGACCATGC 188

45 AARGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAACTGATTTGTGAGG 216

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51159
seq_documentation_block:
ID AAA51159 standard; cDNA; 777 BP.
XX
AC AAA51159;
XX
XX 26-SEP-2000 (first entry)
XX
DE Plant-optimized mutant V. cholerae toxin A subunit R72 coding sequence.
XX
XX
KW Heat-labile toxin; CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
KW adjuvant; anti-bacterial; A72R; ss.
XX
XX Vibrio cholerae.
OS Synthetic.
OS
XX
FH key Location/Qualifiers

FT CDS 1..777
FT /*tag= a
FT /product= mutant_cholera_toxin_subunit_A
XX
XX
XX MO200037609-A2.
XX
XX
XX 29-JUN-2000.
XX
XX
XX
PF 22-DEC-1999; 99WO-US30747.
XX
XX
PR 22-DEC-1998; 98US-0113507.
XX
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX
XX
PI Mason HS, Arntzen CJ;
XX
XX
DR WPI; 2000-442653/38.
DR P-PSDB; AAY96654.
XX
XX
XX
PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX
PS Claim 15; Fig 4C; 103pp; English.
XX
XX
XX This plant-codon optimized cDNA encodes a mutant synthetic
CC Vibrio cholerae cholera toxin (CT) A subunit (CT-A) having a A72R
CC substitution. The sequence has no cryptic signal sequence and no CG
CC (potential methylation sites) sequences. Novel polynucleotides encode a
CC mutant Escherichia coli heat-labile toxin (LT) subunit A (LT-A)
CC polypeptide or a mutant V. cholerae cholera toxin (CT) A subunit (CT-A)
CC polypeptide, which have reduced enzyme activity as compared to the
CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
CC is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible
CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The
CC mutant polypeptides are also useful as adjuvants.
XX
XX
SQ Sequence 777 BP; 210 A; 174 C; 191 G; 202 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAA51159 ..

Align seg 1/1 to: AAA51159 from: 1 to: 777

29 GluYrPheAspArgGlyThrGlnMetAsnIleAsnLeuYrAspHisAl 45
|||||
139 GAGTACTTGTGACAGGGGTACTGACATGCAACCTTTATGACCATGC 188

45 AARGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAACTGATTTGTGAGG 216

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA51160
seq_documentation_block:
ID AAA51160 standard; cDNA; 777 BP.
XX
XX
AC AAA51160;
XX
XX 26-SEP-2000 (first entry)
XX

DE Plant-optimized mutant V. cholerae toxin A subunit G192 coding sequence.
XX
KM Heat-labile toxin: CT-A; CT-B; mutant; transgenic plant; vaccine; oral;
XX adjuvant; anti-bacterial; R192G; ss.
XX
OS Vibrio cholerae.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..777
FT CDS /tag= a
FT /product= mutant_cholera_toxin_subunit_A
XX
PN MO200037609-A2.
PD 29-JUN-2000.
XX
PF 22-DEC-1999; 99WO-US30747.
XX
PR 22-DEC-1998; 98US-0113507.
XX
PA (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX
PI Mason HS, Arntzen CJ;
XX
DR WPI: 2000-442653/38.
DR P-PSDB; AAY96656.
XX
PT New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX
XX
PS Claim 15; Fig 4D; 103pp; English.
XX
XX
CC This plant-codon optimized cDNA encodes a mutant synthetic
CC Vibrio cholerae cholera toxin (CT) A subunit (CT-A) having a R192G
CC substitution. The sequence has no cryptic signal sequence and no CG
CC (potential methylation sites) sequences. Novel polynucleotides encode a
CC mutant Escherichia coli heat-labile toxin (LT) subunit A (LT-A)
CC polypeptide or a mutant V. cholerae cholera toxin (CT) A subunit (CT-A)
CC polypeptide, which have reduced enzyme activity as compared to the
CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
CC is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible
CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae. The
CC mutant polypeptides are also useful as adjuvants.
XX
SQ Sequence 777 BP; 208 A; 174 C; 191 G; 203 T; 1 U; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAS1160 ..

Align seg 1/1 to: AAS1160 from: 1 to: 777

29 GIUTYrPheasPARgLyThrGImeTAsnIleasnLeuTYrAspHISAI 45
|||||
139 GAGTACTTTGACAGGGGACTGACATGCAACCTTTATGACCATGC 188
45 aARGGLyThrGInThrGlyPheValArg 54
|||||
189 AAGGGAACTCAAACTGATTGTGAGG 216

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: AAN30169
seq_documentation_block:
ID AAN30169 standard; DNA; 1148 BP.
XX
XX AAN30169;
XX
AC AAN30169;
XX
DT 25-APR-1992 (first entry)
XX
XX
DE Sequence of PRIT 18014 encoding sub-unit A and B of cholera toxin.
XX
XX Vaccine; cholera; subunit A; subunit B; ss.
XX
XX
OS Vibrio cholerae.
XX
FH Key Location/Qualifiers
FT 1..777
FT misc_feature /tag= a
FT /label= F1
FT /note= "claimed fragment"
FT misc_feature 774..1148
FT /tag= b
FT /label= F2
FT /note= "claimed fragment"
XX
XX
PN EP95452-A.
XX
PD 30-NOV-1983.
XX
XX
PF 23-MAY-1983; 83EP-0091416.
XX
PR 24-MAY-1982; 82US-0381083.
XX
XX
PA (SMIR) SMITH KLINE-RIT.
PI Harford N, Dewilde M;
XX
XX
DR WPI: 1983-834665/49.
XX
PT DNA sequence coding for cholera toxin sub units - for
PT transforming host cells, useful in vaccine prodn.
XX
XX
PS Claim 2; Page 27-28; 46pp; French.
XX
XX
CC The inventors claim a sequence contg. at least one fragment coding
CC for all or part of sub-units A and B of cholera toxin. The entire
CC sequence and two fragments (F1 and F2) are listed in the
CC specification. Also new are recombinant DNA molecules contg. these
CC sequences (spec. PRIT 18014 contg. the whole sequence; PRIT 18041
CC contg. F1 and PRIT 10810 contg. F2) and transformed host cells (spec.
CC E. coli ATCC 39052 contg. the whole sequence, ATCC 39053 contg. F1
CC and ATCC 39051 contg. F2).
XX
SQ Sequence 1148 BP; 382 A; 184 C; 235 G; 347 T; 0 other;

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AAN30169 ..

Align seg 1/1 to: AAN30169 from: 1 to: 1148

29 GIUTYrPheasPARgLyThrGImeTAsnIleasnLeuTYrAspHISAI 45
|||||
139 GAGTACTTTGACAGGGGACTGACATGCAACCTTTATGACCATGC 188
45 aARGGLyThrGInThrGlyPheValArg 54
|||||
189 AAGGGAACTCAAGGATTTGTAGG 216

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AA167270

seq_documentation_block:

ID AA167270 standard; DNA; 2056 BP.

AA167270;

11-FEB-2002 (first entry)

Cholera toxin A and B subunits encoding DNA sequence.

Enterotoxin; mutation; chimeric; immunostimulant; immune response; cholera toxin; ds.

Vibrio cholerae.

Location/Qualifiers

1..550

/*tag= a

551..1327

/*tag= b

/product= "Cholera toxin A subunit"

/note= "AAG65991"

1324..1698

/*tag= c

/product= "Cholera toxin B subunit"

/note= "AAG65992"

/product= "B chain"

1699..2056

/*tag= d

WO200170257-A1.

27-SEP-2001.

16-MAR-2001; 2001WO-US08582.

17-MAR-2000; 2000US-190058P.

(UABR-) UAB RES FOUND.

McGhee J, Kiyono H, Takeda Y, Ohmura M, Yamamoto S;

WPI: 2001-648368/74.

P-SDB; AAG65991, AAG65992.

Chimeric molecule useful as adjuvant for cell-mediated/humoral immunity

comprises first mutated A subunit of first enterotoxin and second

non-mutated subunit from second enterotoxin different from natural

enterotoxin

Disclosure; Fig 4; 23pp; English.

The invention provides a chimeric molecule that comprises a first subunit

which is mutated A subunit of first enterotoxin and a second non-mutated

subunit from a second enterotoxin which is different from the natural

enterotoxin which has been mutated to provide A subunit. A composition of

matter comprising the chimeric molecule in a pharmaceutically acceptable

carrier is useful for obtaining enhanced immune response of an organism

to an antigen by administering it to the antigen. The chimeric molecule

is useful to provide specific immune response to a particular

enterotoxin, as adjuvants for use with unrelated vaccines, and to

customize adjuvants to direct production of cell-mediated or humoral

immune responses. It is also useful as mucosal adjuvant for cell-mediated

or humoral immunity. The present sequence represents the DNA sequence of

Cholera toxin A and B subunits.

Sequence 2056 BP; 603 A; 327 C; 479 G; 647 T; 0 other;

alignment_scores: 26.00 Length: 26

Quality: 26.00

26

26

26

26

26

26

26

26

26

26

26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AA167270

Align seq 1/1 to: AA167270 from: 1 to: 2056

29 GluTrpPheAspArgGlyThrGlnMetAsn1leAsnLeuTyrAspHisAl 45

689 GAGTACTTTCACCGAGGTACTCAATATCATATCACTTATATCATATGC 78

45 AARGGLYThrGlnThrGlyPheValArg 54

739 AAGAGAACTCAGACGCGATTGTAGG 766

seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:ABA92644

seq_documentation_block:

ID ABA92644 standard; DNA; 7586 BP.

ABA92644;

25-MAR-2002 (first entry)

Cholera toxin A1 subunit sequence SEQ ID NO:1.

Cholera toxin; bacterial bleb; vaccine; eukaryotic expression cassette;

gene therapy; ds.

Vibrio cholerae.

WO200189535-A1.

29-NOV-2001.

24-MAY-2001; 2001WO-US16904.

24-MAY-2000; 2000US-206994P.

(UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.

Powell RJ, Hone D;

WPI: 2002-122002/16.

Introducing and expressing a gene in animal cells, useful for

delivering a therapeutic agent to the cells, comprises infecting animal

cells with bacterial blebs containing a eukaryotic expression cassette

encoding the gene

Example 5; Fig 3; 83pp; English.

The present invention describes a method for introducing and expressing

a gene in animal cells, comprising infecting animal cells with bacterial

blebs containing a eukaryotic expression cassette encoding the gene.

The present invention also describes: (1) a bacterial bleb containing a

eukaryotic expression cassette; (2) a transformed animal cell including

a bacterial bleb, or a transfected product of the bacterial bleb;

(3) a method of delivering a eukaryotic expression cassette to an

animal cell, comprising mucosal transfection; and (4) a method of

therapeutically treating an animal with a therapeutic agent encoded by a

polynucleotide coding sequence, by transfecting the animal with a

bacterial bleb containing a eukaryotic expression cassette including the

polynucleotide coding sequence. The method can be used for introducing the

endogenous or foreign genes into animal cells using bacterial blebs as

vectors. The method allows for the delivery of eukaryotic expression

cassettes encoding the endogenous or foreign genes into animal cells or

tissues, and is useful for expressing e.g. vaccine antigens, gene

therapeutic agents, immunoregulatory agents, antisense RNAs, and

catalytic RNAs in animal cells or tissues. The bacterial blebs

containing the eukaryotic expression cassette can also be used to treat

animal cells cultured in vitro. The method can be used in gene therapy.

CC The present sequence represents the cholera toxin A1 subunit sequence,
 CC which is used in an example from the present invention for the production
 CC of non-pyrogenic blebs containing a DNA vaccine.

XX Sequence 7586 BP; 1619 A; 1747 C; 1733 G; 1657 T; 830 other;

alignment_scores:

Quality:	26.00	Length:	26
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x ABA92644 ..

Align seg 1/1 to: ABA92644 from: 1 to: 7586

```

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
87 GAGTACTTACCGAGGTACTCAATGATATTCACCTTATGATCATGC 136
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
137 AAGAGGAACTCAGACGCGATTGTGTAGC 164

```

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAQ12710

seq_documentation_block:
 ID AAQ12710 standard; DNA; 780 BP.

XX AAQ12710;

DT 08-OCT-1991 (first entry)

XX Cholera toxin A1 fragment.

XX Hybrid: fusion: membrane translocation; binding region: HIV;

KM infection: toxin: steroid; hormone: monoclonal antibody; antigen;

KM diphtheria; exotoxin; phenylketonuria; cholera; interleukin; IL-2;

KM protease; epidermal growth factor; ricin; tetanus; hexosaminidase;

XX Shiga-like toxin A; ligand; insulin; nuclease; ss.

OS Vibrio cholerae.

XX key

FT CDS

FT 1..774

FT /*tag= a

FT /label= CT-A1

XX WO9109671-A.

XX 11-JUL-1991.

XX 21-DEC-1990;

XX 90WO-US07619.

XX 14-JUN-1990;

XX 90US-0538276.

XX 22-DEC-1989;

XX 89US-0456095.

XX (SERA-) SERAGEN INC.

XX Murphy JR;

XX WPI: 1991-222845/30.

XX P-PSDB; AARI3117.

XX Hybrid molecules for targeting chemical entity to cell - have

XX membrane trans-locating and cell binding-regions and used to

XX treat HIV infection, genetic enzyme-deficiency disorders etc.

XX Disclosure: Fig 5(1-2); 59pp; English.

XX Hybrid molecules are produced by covalently linking

XX (1) a portion (A) of the binding domain of a cell-binding ligand,

CC allowing binding of the mol. to an animal cell;

CC (2) a portion (B) of a translocation domain of a protein able to

CC translocate (C) across the cell cytoplasmic membrane, and the cell.

CC and (3) a portion (C) which is to be introduced into the cell.

CC (A) is derived from a steroid or polypeptide hormone, a single-chain

CC analogue of a monoclonal antibody able to bind an antigen expressed

CC on the cell surface, or a polypeptide toxin.

CC (B) is derived from a toxin (e.g. diphtheria toxin or Pseudomonas

CC exotoxin A).

CC (A) may be derived from insulin, interleukins 2, 3 or 6 or

CC epidermal growth factor.

CC Sulfable enzymes in (C) include cholera toxin (encoded by this

CC sequence), ricin, tetanus toxin, hexosaminidase A, protease,

CC nuclease, etc.

CC Specified examples are CT-A/DT-B'/IL-2, SLTA/DT-B'/IL-2,

CC ricin A/DT-B'/IL-2, HIV-BP/DT-B'/IL-2 and the phenylalanine

CC hydroxylase-DT-B' or their biologically active mutants.

CC (CT-A= cholera toxin, DT-B'= truncated diphtheria toxin,

CC SLTA= Shiga-like toxin A; HIV-BP= HIV protease binding protein.

CC See also AAQ12710-12.

XX Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other;

XX alignment_scores:

Quality:	25.00	Length:	25
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-528-682-1 x AAQ12710 ..

Align seg 1/1 to: AAQ12710 from: 1 to: 780

```

30 TyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaAr 46
|||||
142 TACTTTGACCGAGGTACTCAATGATATTCACCTTATGATCATGCAG 191
46 gGlyThrGlnThrGlyPheValArg 54
|||||
192 AGGAACCTCAGACGCGATTGTGTAGC 216

```

seq_name: /SID5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAI91636

seq_documentation_block:

ID AAT91636 standard; DNA; 780 BP.

XX AAT91636;

XX 27-MAR-1998 (first entry)

XX Vibrio cholerae toxin gene.

XX Cholera toxin A1; hybrid protein; cell delivery;

KM cell binding ligand; translocation domain; diphtheria toxin B';

KM interleukin-2; leukaemia; therapy; ss.

XX Vibrio cholerae.

XX key

FT CDS

FT 1..777

FT /*tag= a

FT /transl_except= (pos:358..360, aa:Ala)

XX US5668255-A.

XX 16-SEP-1997.

XX 07-JUN-1984;

XX 84US-0618199.

XX 27-JUN-1991;

XX 91US-0722484.

XX 07-JUN-1984;

XX 84US-0618199.


```

PR 25-APR-1985; 85US-0726808.
PR 07-JUN-1985; 85US-0742554.
PR 22-DEC-1989; 89US-0456095.
PR 14-JUN-1990; 90US-0538276.
PR 04-AUG-1993; 93US-0102387.
XX
XX (SERA-) SERAGEN INC.
XX
XX Murphy JR;
XX
XX WPI: 1997-470103/43.
XX
XX P-PSDB; AAM25785.
XX
XX New hybrid molecules for delivery of agents to cells - comprise a
XX binding domain of a cell binding ligand and a portion of a
XX translocation domain of a protein
XX
XX Example 2; Fig 5A-B; 30pp; English.
XX
XX This sequence comprises the Vibrio cholerae gene that encodes
XX cholera toxin (see AAM25785). DNA encoding an enzymatically active
XX A1 fragment of the cholera toxin was used to construct a cholera
XX toxin A1-diphtheria toxin B'-interleukin-2 (CTA-DTB-IL2) gene that
XX was expressed in E. coli. The hybrid protein can be isolated and
XX used e.g. as an adjunct to treatment with DTB-IL2 hybrid, which
XX targets the IL2 receptor, such as certain leukaemic T cells.
XX bearing the IL2 receptor, such as certain leukaemic T cells.
XX Prior application of the CTA-DTB-IL2 hybrid alleviates the
XX problem of endogenous IL2 competing with DTB-IL2 hybrid for IL2
XX receptors on T cells. Claimed hybrid proteins comprise a
XX translocation domain and a cell binding domain from e.g. a hormone,
XX growth factor or polypeptide toxin. The hybrid molecules can be
XX used for the delivery of agents (e.g. therapeutic genes, toxins,
XX detectable labels) into cells. The use of a translocation
XX mechanism ensures that the hybrid will be effective in relatively
XX low doses, since a high proportion of the substance of interest
XX will be taken into the targeted cells. The hybrid molecules can be
XX manufactured as a single hybrid recombinant protein, permitting
XX reproducibility, consistency, and the precise control of
XX composition.
XX
XX Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:
XX
XX alignment_scores:
XX      Quality: 25.00      Length: 25
XX      Ratio: 1.000      Gaps: 0
XX      Percent Similarity: 100.000      Percent Identity: 100.000
XX
XX alignment_block:
XX US-09-528-682-1 x AAT91636 ..
XX
XX Align seg 1/1 to: AAT91636 from: 1 to: 780
XX
XX 30 TyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaIar 46
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 142 TACTTTGCGCAGGCTACTCAATGATATCAACCTTATGATCATGCAAG 191
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 46 GGLYThrGlnThrGlyPheValArg 54
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 192 AGGAACCTCAGACGGGATTTTGTAGG 216
XX
XX seq_name: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAZ30661
XX seq_documentation_block:
XX ID AAZ30661 standard; DNA: 780 BP.
XX
XX AC AAZ30661;
XX
XX XX
XX DT 15-FEB-2000 (first entry)
XX
XX XX
XX DE Vibrio cholerae A1 fragment toxin gene.
XX

```

```

KW Recombinant; hybrid; binding domain; ligand; animal cell; diphtheria;
KW translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;
KW shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;
KW cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison;
KW adipocyte; cancer; virus; infection; antibody; ss.
XX
XX Vibrio cholerae.
XX
XX US5965406-A.
XX
XX 12-OCT-1999.
XX
XX 07-JUN-1995; 95US-0488246.
XX
XX 04-AUG-1993; 93US-0102387.
XX 07-JUN-1984; 84US-0618199.
XX 27-JUN-1991; 91US-07222484.
XX 25-APR-1985; 85US-0726808.
XX 07-JUN-1985; 85US-0742554.
XX 22-DEC-1989; 89US-0456095.
XX 14-JUN-1990; 90US-0538276.
XX
XX (SERA-) SERAGEN INC.
XX
XX Murphy JR;
XX
XX WPI: 1999-632431/54.
XX P-PSDB; AAY55690.
XX
XX Recombinant DNA molecule encoding a three part hybrid protein used in
XX the treatment of Aids and genetic deficiency diseases -
XX
XX Example 2; Fig 5; 31pp; English.
XX
XX The invention relates to a recombinant DNA molecule encoding a hybrid
XX protein comprising three parts: (a) the first part comprises a portion
XX of the binding domain of a cell-binding polypeptide ligand allowing the
XX hybrid protein to bind to an animal cell; (b) the second part comprises
XX a portion of a translocation domain of a naturally occurring protein
XX selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera
XX toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus
XX toxin, which translocate the third part of the across the cytoplasmic
XX membrane into the cytosol of the cell; and (c) the third part comprises
XX a polypeptide entity to be introduced into the cell, which is non-native
XX to the naturally occurring protein of (b). This sequence represents the
XX Vibrio cholerae toxin A1 fragment gene for use in generating the hybrid
XX of the invention. The hybrid molecule enables the direction of
XX appropriate therapy to affected cells, allowing them to function properly
XX and alleviate or cure the disease. The hybrid is especially used in
XX treating genetic deficiency diseases, by delivering to affected cells
XX an enzyme supplying the missing function, to supplementing cellular
XX levels of a particular enzyme or a scarce precursor or cofactor, to
XX directing toxins or other poisons to destroy particular cells (such as
XX adipocytes, cancer cell, or virus infected-cells), to counteracting viral
XX infections such as HIV, by introducing appropriate antibodies to viral
XX proteins. It is also involved in the process of getting non-therapeutic
XX substances such as detectable labels into cells.
XX
XX Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other:
XX
XX alignment_scores:
XX      Quality: 25.00      Length: 25
XX      Ratio: 1.000      Gaps: 0
XX      Percent Similarity: 100.000      Percent Identity: 100.000
XX
XX alignment_block:
XX US-09-528-682-1 x AAZ30661 ..
XX
XX Align seg 1/1 to: AAZ30661 from: 1 to: 780
XX
XX 30 TyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaIar 46
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||

```


142 TACTTTGACCGAGTCAATGATATCAACCTTATGATCATGCAG 191
 46 ggGlyThrGlnThrGlyPheValArg 54
 192 AGGAACTCAGACGCGGATTTGTTAGG 216

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA2000.DAT:AA290017

seq_documentation_block:
 ID AA290017 standard; DNA; 780 BP.

AA290017:

05-MAY-2000 (first entry)

Vibrio cholerae toxin gene.

Cholera toxin; hybrid protein; translocation domain; cell binding domain;
 genetic deficiency disease; cell targeting; cell destruction; cancer;
 adipocyte; enzyme delivery; anti-viral; HIV; ds.

Vibrio cholerae.

US6022950-A.

08-FEB-2000.

07-JUN-1995; 95US-0479510.

07-JUN-1984; 84US-0618159.

27-JUN-1991; 91US-0722484.

25-APR-1985; 85US-0726808.

07-JUN-1985; 85US-0742554.

22-DEC-1989; 89US-0456095.

14-JUN-1990; 90US-0538276.

04-AUG-1993; 93US-0102387.

(SERA-) SERAGEN INC.

Murphy JR;

WPI; 2000-160390/14.

P-PSDB; AAY78590.

Example 2: Fig 5: 32pp; English.

This sequence represents the Vibrio cholerae toxin gene. The toxin can be included in the hybrid protein of the invention and used to destroy or modify the cell that the hybrid protein is targeted to. The hybrid protein comprises a first part which is a portion of the binding domain of a cell-binding ligand, effective to cause the hybrid molecule to bind to a cell of an animal. The second part comprises a portion of a translocation domain of a naturally occurring protein (e.g. the translocation domain of diphtheria toxin) the second part translocates the third part across the cytoplasmic membrane and into the cytosol of the cell. The third part comprises a chemical entity to be introduced into the cell, where each of the first and third part is non-native with respect to naturally occurring protein, and the covalent bond attaching the second and third part is cleavable. The toxin encoded by the present sequence can form part of the third portion of the hybrid protein. The cell binding domain binds to a specific cell and the translocation domain transfers the hybrid molecule across the cell membrane into the cytosol. The third part of the protein, linked to the translocation domain through a cleavable bond, can then carry out its function. The hybrid molecules are useful for treating genetic deficiency diseases by delivering to affected cells an enzyme supplying the missing function, to supplement cellular levels of a particular enzyme or a scarce precursor or cofactor, to direct toxins or other poisons to destroy particular cells (such as adipocytes, cancer cells, or virus-infected cells), and to counteract

CC viral infections such as HIV by introducing into appropriate cells
 CC antibodies to viral proteins.

Sequence 780 BP; 243 A; 130 C; 169 G; 238 T; 0 other;

Alignment_scores:

Quality:	Ratio:	Length:	Gaps:
25.00	1.000	25	0
Percent Similarity: 100.000	Percent Identity: 100.000		

alignment_block:

US-09-528-682-1 x AA290017

Align seg 1/1 to: AA290017 from: 1 to: 780

30 TyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAlaAr 46
 142 TACTTTGACCGAGTCAATGATATCAACCTTATGATCATGCAG 191
 46 ggGlyThrGlnThrGlyPheValArg 54
 192 AGGAACTCAGACGCGGATTTGTTAGG 216

seq_name: /SID55/gcdata/geneseq/geneseqn-emb1/NA1993.DAT:AA051327

seq_documentation_block:

ID AA051327 standard; DNA; 723 BP.

AA051327:

08-DEC-1993 (first entry)

Encodes Asp-53 cholera toxin subunit A.

Immunogenic detoxified CT-A; cholera vaccine; adenylate cyclase;

protomer A; site-directed mutagenesis; reduced toxicity;

ADP-ribosyltransferase activity; ss.

Vibrio cholerae.

Key Location/Qualifiers

misc_difference 156..158

W09313202-A.

08-JUL-1993.

30-DEC-1992; 92WO-EP03016.

31-DEC-1991; 91IT-0M13513.

(BIOC-) BIOCIINE SCLAVO SPA.

Domenighini M, Hol W, Piazza M, Rappuoli R;

WPI; 1993-227320/28.

P-PSDB; AAR44026.

Immunogenic detoxified mutant cholera toxin and heat labile toxin

- useful as vaccines against infection by Vibrio cholerae and

enterotoxin producing Escherichia coli

Claim 3: Fig 2 and Page 48; 60pp; English.

The wild-type sequence coding for the A subunit of the cholera toxin (CT-A) (Mekalanos et al., Nature 306, 551 (1983) - see AA046318) was subjected to site-directed mutagenesis. Certain mutations were found to reduce toxicity (see AA051327-051334). The invention relates to immunogenic, detoxified CT-A proteins and their use in vaccines to protect against Vibrio cholerae. Sequence AA051327 is a

CC combination of the wild-type coding sequence and the mutagenic
 CC primer sequence used to introduce the preferred mutation.
 XX
 SQ Sequence 723 BP; 228 A; 126 C; 163 G; 206 T; 0 other;

alignment_scores:
 Quality: 24.00 Length: 24
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AAQ31263 ..

Align seg 1/1 to: AAQ31263 from: 1 to: 723

29 GluTyrPheAspArgGlyThrGlnMetAsnLeuTyrAspHisAl 45
 ||||||||||||||||||||||||||||||||||||||||
 85 GAGTACTTGTGACGAGTACTCAATGATATACCTTATGATCATGC 134
 45 AARGlyThrGlnThrGlyPhe 52
 ||||||||||||||||||||||||
 135 AAGAGGAACTCAGACGCGATT 156

seq_name: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAQ31263

seq_documentation_block:

ID AAQ31263 standard: DNA; 777 BP.

XX AAQ31263;

XX 25-MAR-1993 (first entry)

XX Mutant A subunit of CT.

XX Cholera toxin; A1; A2; B; subunit; mutation; ss.

XX Vibrio cholerae.

XX Key Location/Qualifiers

FT CDS 1..777

FT sig_peptide 1..34

FT mat_peptide 55..774

FT misc_feature 630

FT misc_feature 636

FT misc_feature 637..777

FT misc_feature 637..777

FT misc_difference 73..75

FT misc_difference 79..81

FT misc_difference 85..87

FT misc_difference 184..186

FT misc_difference 262..264

FT misc_difference 388..390

FT /tag= m
 FT /note= "Glu112 codon mutation, see CC"
 FT misc_feature 588
 FT /tag= n
 FT /note= "start of C-terminal truncation, see CC"

XX WO9219265-A.
 XX 12-NOV-1992.

XX 04-MAY-1992; 92MO-US03703.

XX 02-MAY-1991; 91US-0694733.

XX (AMGE-) AMGEN.
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX Burnette WN, Kaslow HR;

XX WPI, 1992-398532/48.

XX P-PSDB; AAR28830.

XX Recombinant DNA encoding analogue of Vibrio cholerae catalytic
 PT sub-unit - useful as a vaccine for preventing cholera, has little
 PT or no catalytic activity

XX

XX Claim 7-8; Page 49-50 + Fig 1A; 74pp; English.

XX A recombinant mol. encoding an analogue of the catalytic subunit of
 CC cholera toxin (CT), where the analogue has reduced or no catalytic
 CC activity associated with CT. reactogenically is an analogue of the A
 CC region, esp. A1 region, comprising a site-specific mutation in the
 CC region bounded by the codons for Met1 and Arg192 or Ser194, esp. one
 CC or more of Arg7, Arg11, Asp9, His44, His70 and Glu112 (see tag h-m
 CC of Features Table). The analogue may also comprise a truncation of
 CC the C-terminal portion starting at Trp179 (see tag n of Features Table).
 CC The DNA may also encode subunit B of CT, esp. the B oligomer in
 CC native form or which has been genetically engineered (see AAQ31264).
 XX

SQ Sequence 777 BP; 235 A; 125 C; 165 G; 234 T; 18 other;

alignment_scores:
 Quality: 22.00 Length: 22
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AAQ31263 ..

Align seg 1/1 to: AAQ31263 from: 1 to: 777

81 SerThrTyrTrpIleValIleAlaThrAlaProAsnMetPheAsn 97
 ||||||||||||||||||||||||||||||||||||||||
 295 TCTACTTATTATATATATGTTATAGCCACTGACCAACATGTTAAAG 344
 97 IAsnAspValLeuGly 102
 ||||||||||||||||
 345 TAATGATGATATAGCG 360

gb_gss:BH268086	+	8.00	124.52	1.1e+03	539	BH268086	CH230-56A10.TJ	CHORI-2	gb_gss:BH249846	-	8.00	121.70	1.6e+03	764	BH249846	BOGAM63TF	BOGA	Bras	
gb_est1:AV882538	+	8.00	124.51	1.1e+03	540	AV882538	AV882538	Noti	Satoh un	gb_est2:BI658656	-	8.00	121.66	1.6e+03	768	BI658656	60328613F1	NIH_MGC	
gb_est1:BM134967	+	8.00	124.46	1.1e+03	543	BM134967	WMHO0460.C12.E23S	Whee	gb_est2:BG446666	-	8.00	121.52	1.6e+03	781	BG446666	602420826F1	NIH_MGC		
gb_gss:BA2159324	+	8.00	124.45	1.1e+03	543	AL479551	Ts_burice1	Shepard	genc	gb_est2:BG789155	-	8.00	121.51	1.6e+03	782	BG789155	SPBA000911F	Seei	
gb_gss:BF26615	+	8.00	124.43	1.1e+03	545	AO179324	Hs_3165.A2	Al12-77	CIT	gb_gss:AO683732	+	8.00	121.44	1.6e+03	789	AO683732	nxb0070411F	CUGI	
gb_est1:AV755146	+	8.00	124.35	1.1e+03	551	AV727316	Clona	Intestinalis	get	gb_est1:BE033932	+	8.00	121.43	1.6e+03	790	BE033932	M602C06	MG	
gb_gss:BM338148	+	8.00	124.30	1.1e+03	554	BM328148	WMHO0945.F06.K12S	Whee	gb_gss:BM463146	-	8.00	121.27	1.7e+03	805	BM463146	BOGYO44TF	BOGY		
gb_gss:BM338148	+	8.00	124.30	1.1e+03	554	BM328148	CH230-12412.TVB	CHORI-	gb_est1:AV6611452	-	8.00	121.25	1.7e+03	807	AV6611452	AV681462	GA		
gb_est1:AL510917	+	8.00	124.24	1.1e+03	558	AL510917	AL510917	Hordeum	val19g	gb_gss:AO745347	-	8.00	121.17	1.7e+03	816	AO745347	HS_2276	AL	
gb_est2:BM178462	+	8.00	124.21	1.1e+03	560	BM178462	sa773502.y1	Gm-c1072	G	gb_est2:BG369434	-	8.00	121.15	1.7e+03	818	BG369434	HVSMEM0024G08F	Hor	
gb_gss:AZ709793	+	8.00	124.16	1.1e+03	564	AZ709793	RPCT-24-116118.TJ	RPCT	gb_est2:BI256096	+	8.00	121.14	1.7e+03	819	BI256096	602978931F1	NCL_CGA		
gb_est1:AA739702	+	8.00	124.13	1.1e+03	566	AA739702	467	PIF62	Pinus	teda	gb_est2:BI956846	-	8.00	121.14	1.7e+03	819	BI956846	HVSMEM0005K09F	Hor
gb_est2:BG414316	+	8.00	124.13	1.1e+03	566	BG414316	HVSMEM0001G05F	Hordeum	gb_gss:CNS02CA6	-	8.00	121.05	1.7e+03	828	AL199887	Tetradodon	nigrovitr		
gb_est1:AV033568	+	8.00	124.06	1.2e+03	571	AV933568	AV933568	K. Sato	unpnt	gb_gss:CNS03AR2	-	8.00	121.04	1.7e+03	829	AL264335	Tetradodon	nigrovitr	
gb_est2:BG414329	+	8.00	124.06	1.2e+03	571	BG414329	HVSMEM0001E09F	Hordeum	gb_est1:BE194407	-	8.00	120.89	1.7e+03	844	BE194407	HVSMEM0085H01F	Hor		
gb_est2:BR483656	+	8.00	124.00	1.2e+03	575	BF483656	WHE2336.C12.F24ZS	Whee	gb_est2:BM018143	-	8.00	120.86	1.7e+03	847	BM018143	603645733F1	NIH_MGC		
gb_est2:BF065521	+	8.00	123.97	1.2e+03	577	BF065521	HV_CEB0012K16F	Hordeum	gb_gss:CNS04S0U	-	8.00	120.80	1.8e+03	854	AL305679	Tetradodon	nigrovitr		
gb_est1:AL510920	-	8.00	123.96	1.2e+03	578	AL510920	AL510920	Hordeum	val19g	gb_gss:CNS04S0U	-	8.00	120.76	1.8e+03	858	BO8256	F27A9-Sp6.1	IGF	
gb_gss:AO2313446	+	8.00	123.90	1.2e+03	582	AO313446	RPCT11-101B1.TJ	RPCT-1	gb_est2:BF575094	-	8.00	120.73	1.8e+03	861	BF575094	602134766F1	NIH_MGC		
gb_gss:AO2314058	+	8.00	123.87	1.2e+03	584	AO274058	nxb00032118F	CUGI	R16c	gb_est2:BF120183	-	8.00	120.65	1.8e+03	870	BF120183	601756489F1	NCL_CGA	
gb_est2:BF640414	+	8.00	123.83	1.2e+03	587	BF640414	SD3280.5	Prime	SD	Dros	gb_est2:BF575766	-	8.00	120.52	1.8e+03	884	BF575766	602135396F1	NIH_MGC
gb_est1:AM014756	+	8.00	123.82	1.2e+03	588	AM014756	UT-H-B10-aae-e-11-0-01			gb_est2:BI222903	+	8.00	120.47	1.8e+03	889	BI222903	602941152F1	NIH_MGC	
gb_gss:BM260927	+	8.00	123.74	1.2e+03	594	BM260927	CH230-90H24.TV	CHORI-2	gb_gss:CNS04HK9	-	8.00	120.44	1.9e+03	917	AL291042	Tetradodon	nigrovitr		
gb_est1:AV036810	-	8.00	123.68	1.2e+03	598	AV936810	AV936810	K. Sato	unpnt	gb_gss:CNS02FC6	-	8.00	120.22	1.9e+03	949	AL179456	Tetradodon	nigrovitr	
gb_est1:AV036810	-	8.00	123.70	1.2e+03	598	BI074785	IP1_15-B03.B1	AO02	Imm	gb_gss:CNS03AR2	-	8.00	119.94	2.0e+03	959	AL334815	Tetradodon	nigrovitr	
gb_est1:AV037228	-	8.00	123.64	1.2e+03	601	AV937228	AV937228	K. Sato	unpnt	gb_gss:CNS03AR2	-	8.00	119.86	2.0e+03	959	AL304362	Tetradodon	nigrovitr	
gb_est2:BI074770	+	8.00	123.64	1.2e+03	601	BI074770	IP1_15-C06.B1	AO02	Imm	gb_est2:BI597916	-	8.00	118.39	2.4e+03	1149	BI597916	603061952F1	NIH_MGC	
gb_gss:AO116267	+	8.00	123.63	1.2e+03	602	AO116267	RPCT11-42M4.TK-1	RPCT-		gb_gss:AO636471	-	8.00	119.43	2.1e+03	1011	AG103651	Pan	troglydtes	
gb_gss:AZ993578	+	8.00	123.57	1.2e+03	606	AZ993578	2M027BK1OR	Mouse	10kb	gb_est2:BF827346	-	8.00	119.30	2.1e+03	1027	BF827346	602101178F1	NCL_CG	
gb_est1:BI421028	+	8.00	123.56	1.2e+03	607	BI421028	LJNES6413R	Lotus	japc	gb_est2:BF827346	-	8.00	119.22	2.2e+03	1038	BE477107	PSR6421	ITRC	
gb_gss:AO276759	-	8.00	123.52	1.2e+03	610	AO276759	CITB1-E1-2512K4.TR	CIT		gb_gss:AO041888	-	8.00	119.17	2.2e+03	1044	BI438631	Tetradodon	nigrovitr	
gb_gss:AO276759	-	8.00	123.52	1.2e+03	610	AO276759	nxb0003161F	CUGI	R16c	gb_est1:AA771007	-	8.00	119.17	2.2e+03	1081	BI438631	Tetradodon	nigrovitr	
gb_gss:AO276759	-	8.00	123.39	1.3e+03	620	BE291294	WHE2221.E05.109ZS	AGS1		gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_gss:AO276759	-	8.00	123.34	1.3e+03	624	AO290882	RPCT-23-458C16.TJ	RPCT		gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM076951	fj33a01.x1	zebrafish	a	gb_est1:AA771007	-	7.00	122.62	1.4e+03	81	AA771007	602991853F1	NCL_CG	
gb_est1:AM076951	+	8.00	123.48	1.2e+03	613	AM0													

gb_gss:BH081365	+	7.00	116.38	3.1e+03	175	BH081365	RPCT-24-292N3.TJ	RPCT-24-292N3.TJ	gb_est2:BH040696	+	7.00	114.41	4.0e+03	223	B1040696	CM4-NT0287-120201-8
gb_est2:BF659767	-	7.00	116.33	3.1e+03	176	BF659767	ma31c07.y1	NCI CGAP	gb_est1:AV280455	+	7.00	114.34	4.0e+03	225	AV280455	AV280455
gb_gss:AH355670	-	7.00	116.19	3.2e+03	179	BH266086	CH230-136D12.TV	CHOR1	gb_est1:AV2898791	+	7.00	114.34	4.0e+03	225	AV2898791	AV2898791
gb_est1:AM242371	+	7.00	116.10	3.2e+03	181	AV555670	AV555670	Arabidopsis thaliana	gb_est1:AV308032	-	7.00	114.31	4.0e+03	226	AV308032	m160c11.r1 Soares m
gb_est1:AM242371	+	7.00	116.06	3.2e+03	182	AM242371	u028C04.x1	NCI CGAP	gb_est1:AV308032	-	7.00	114.31	4.0e+03	226	AV308032	m160c11.r1 Soares m
gb_est2:BG823211	+	7.00	115.02	3.2e+03	183	AM70344	vq14906.r1	Bartsch st	gb_est1:BB201443	+	7.00	114.31	4.0e+03	226	BB201443	BB201443
gb_est1:AV039349	+	7.00	115.97	3.3e+03	184	AV039349	PM1-CN0098-150201-011-	Mus musculus	gb_est1:BB201443	+	7.00	114.31	4.0e+03	226	BB201443	BB201443
gb_est1:BB046590	+	7.00	115.97	3.3e+03	184	BB046590	BB046590	Riken full-16	gb_est1:BB195952	+	7.00	114.24	4.1e+03	228	BB195952	BB195952
gb_est1:AV246346	+	7.00	115.93	3.3e+03	185	AV246346	AV246346	Riken full-16	gb_gss:AV359348	+	7.00	114.24	4.1e+03	228	AV359348	IM0102609P Mouse 10
gb_est1:AM874145	+	7.00	115.88	3.3e+03	186	AM874145	hg85C08.x1	NCI CGAP	gb_gss:AV359348	+	7.00	114.24	4.1e+03	228	AV359348	IM0102609P Mouse 10
gb_est2:BI189294	+	7.00	115.88	3.3e+03	186	BI189294	ez6065.r1	Fusarium sf	gb_est1:AM330151	+	7.00	114.16	4.1e+03	230	AM330151	RC3-S70281-121299-0
gb_est1:BI189294	+	7.00	115.88	3.3e+03	186	BI189294	ez6065.r1	Fusarium sf	gb_est1:AM330151	+	7.00	114.16	4.1e+03	230	AM330151	RC3-S70281-121299-0
gb_est1:BB060839	-	7.00	115.80	3.3e+03	188	BB060839	BB060839	Riken full-16	gb_est1:AV1461091	+	7.00	114.09	4.2e+03	232	AV1461091	UI-R-B50-9m-w-f-01-0
gb_est1:AV092156	+	7.00	115.75	3.4e+03	189	AV092156	AV092156	Mus musculus	gb_est1:AA755409	+	7.00	114.09	4.2e+03	232	AA755409	m73606.r1 Knowles m
gb_est2:BI052428	+	7.00	115.75	3.4e+03	189	BI052428	PMO-GN0210-200101-002-	Arabidopsis thaliana	gb_est1:AV322182	+	7.00	114.09	4.2e+03	232	AV322182	AV322182
gb_est1:BE520222	+	7.00	115.75	3.4e+03	189	BE520222	M1A2XTM	Arabidopsis thaliana	gb_est1:BB150835	+	7.00	114.09	4.2e+03	232	BB150835	BB150835
gb_est1:AA230954	+	7.00	115.71	3.4e+03	190	AA230954	mw15f09.r1	Soares mous	gb_est1:BB205054	+	7.00	114.06	4.2e+03	233	BB205054	BB205054
gb_est2:BI060066	-	7.00	115.71	3.4e+03	190	BI060066	IL3-UN0114-160101-419-	Arabidopsis thaliana	gb_est1:AV243330	+	7.00	114.02	4.2e+03	234	AV243330	AV243330
gb_est1:AV1787052	-	7.00	115.67	3.4e+03	191	AV1787052	k860781.r1	Aspergillus	gb_est1:AM534811	+	7.00	114.02	4.2e+03	234	AM534811	UI-R-B50-9m-w-f-01-0
gb_gss:BH36347	-	7.00	115.67	3.4e+03	191	BH36347	AG-ND-142021.TF	ND-7M	gb_est1:BB306526	+	7.00	114.02	4.2e+03	234	BB306526	BB306526
gb_est1:AM903803	+	7.00	115.63	3.4e+03	192	AM903803	CM4-NN1037-200400-152-	Arabidopsis thaliana	gb_est1:AV320137	+	7.00	114.02	4.2e+03	234	AV320137	F30137
gb_est1:AV242562	+	7.00	115.59	3.4e+03	193	AV242562	AV242562	Medaka ovary	gb_est1:BB139780	+	7.00	113.99	4.2e+03	235	BB139780	BB139780
gb_est2:BI012349	+	7.00	115.54	3.4e+03	194	BI012349	PM1-GN0065-050301-017-	Arabidopsis thaliana	gb_est1:BB139780	+	7.00	113.99	4.2e+03	235	BB139780	BB139780
gb_gss:BH267920	-	7.00	115.46	3.5e+03	196	BH267920	CH230-21564.TV	CHOR1-2	gb_est1:AV270521	+	7.00	113.96	4.2e+03	236	AV270521	AV270521
gb_gss:BH078864	-	7.00	115.30	3.6e+03	200	BH078864	RPCT-24-292D17.TV	RPCT	gb_est1:AM984505	+	7.00	113.96	4.2e+03	236	AM984505	PM1-HN0012-220300-0
gb_est1:AV284282	+	7.00	115.26	3.6e+03	201	AV284282	AV284282	Riken full-16	gb_est1:AM984505	+	7.00	113.96	4.2e+03	236	AM984505	PM1-HN0012-220300-0
gb_gss:AM889620	-	7.00	115.26	3.6e+03	201	AM889620	RPCT-24-168E11.TV	RPCT	gb_est1:AM984505	+	7.00	113.96	4.2e+03	236	AM984505	PM1-HN0012-220300-0
gb_gss:AM889620	-	7.00	115.22	3.6e+03	201	AM889620	RPCT-24-168E11.TV	RPCT	gb_est1:AM984505	+	7.00	113.96	4.2e+03	236	AM984505	PM1-HN0012-220300-0
gb_gss:AM889620	-	7.00	115.22	3.6e+03	201	AM889620	RPCT-24-168E11.TV	RPCT	gb_est1:AM984505	+	7.00	113.96	4.2e+03	236	AM984505	PM1-HN0012-220300-0
gb_est1:BB173809	+	7.00	115.18	3.6e+03	203	BB173809	BB173809	Riken full-16	gb_est2:BB162369	+	7.00	113.89	4.3e+03	238	BB162369	BB162369
gb_est1:AA351870	+	7.00	115.14	3.6e+03	204	AA351870	ESY9743	Infant brain	gb_est1:AA250373	+	7.00	113.85	4.3e+03	239	AA250373	mw16b08.r1 Soares m
gb_gss:BH507645	-	7.00	115.10	3.7e+03	205	BH507645	B0GR11TR	B0GR brassica	gb_est1:BB199004	+	7.00	113.85	4.3e+03	239	BB199004	BB199004
gb_est2:BG558458	-	7.00	115.06	3.7e+03	206	BG558458	RH122.67	G09.y1	AA03	+	7.00	113.85	4.3e+03	239	AA03	PMO-H70333-080500-0
gb_gss:BM232724	-	7.00	115.02	3.7e+03	207	BM232724	CH230-173M23.TV	CHOR1-2	gb_gss:AM721052	+	7.00	113.85	4.3e+03	239	AM721052	RPCT-24-156D3.TV
gb_est1:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV
gb_est2:BM404638	+	7.00	115.02	3.7e+03	207	BM404638	ESY578965	potato roots	gb_est1:AM882074	+	7.00	113.85	4.3e+03	239	AM882074	RPCT-24-156D3.TV

gb_gss:AZ241407	+	7.00	113.46	4.5e+03	251	AZ241407	RPCT-23-89H4.TJ	RPCT-2	gb_estt1:BB353901	-	7.00	112.60	5.0e+03	279	BB353901	BB353901	RIKEN	full		
gb_gss:BBH31295	+	7.00	113.46	4.5e+03	251	BBH31295	CH230-53E18.TJ	CHORI-2	gb_estt2:BF33135	+	7.00	112.60	5.0e+03	279	BF33135	EST108835	Rat PC-12	C		
gb_estt1:AZ25878	+	7.00	113.42	4.5e+03	252	AZ25878	a125e01.s1	Soares	test	gb_estt1:BF33135	+	7.00	112.60	5.0e+03	279	BF33135	EST108835	Rat PC-12	C	
gb_estt1:AV275013	-	7.00	113.42	4.5e+03	252	AV275013	AV275013	RIKEN	full-16	gb_estt1:AV074256	-	7.00	112.57	5.0e+03	280	AV074256	AV074256	Mus	muscul	
gb_estt1:BE183175	-	7.00	113.36	4.5e+03	253	BE183175	RC4-H10664-250400-012-			gb_estt1:BE004191	-	7.00	112.57	5.0e+03	280	BE004191	CNO-RN0103-180300-2			
gb_estt1:AW199563	-	7.00	113.36	4.6e+03	254	AW199563	PM2-DW0053-030300-003-			gb_estt1:AV057774	-	7.00	112.54	5.1e+03	281	AV057774	AV057774	Mus	muscul	
gb_estt2:BI3979796	+	7.00	113.36	4.6e+03	254	BI3979796	NKPV_107_F05.F	NXPV (N		gb_estt1:BB434311	-	7.00	112.54	5.1e+03	281	BB434311	BB434311	RIKEN	full	
gb_estt2:BI74827	+	7.00	113.36	4.6e+03	254	BI74827	81	Deletion-treated	Bra	gb_estt2:BB435255	-	7.00	112.54	5.1e+03	281	BB435255	1K13H3.abi	Bos	tau	
gb_gss:BF12025	+	7.00	113.33	4.6e+03	255	BF12025	HSPD14762	H3	Hom	SP16	gb_estt1:AV036547	-	7.00	112.51	5.1e+03	282	AV036547	AV036547	Richt	hemt
gb_gss:BBH122334	+	7.00	113.33	4.6e+03	255	BBH122334	RPCT-24-273B23.TJ	RPCT		gb_estt1:AA358866	-	7.00	112.46	5.1e+03	284	AA358866	EST65567	Richt	hemt	
gb_estt1:AV055477	-	7.00	113.30	4.6e+03	256	AV055477	AV054777	Mus	muscul	gb_estt1:BB230078	-	7.00	112.51	5.1e+03	282	BB230078	BB230078	RIKEN	full	
gb_estt1:AV125212	-	7.00	113.30	4.6e+03	256	AV125212	AV125212	Mus	muscul	gb_estt1:BB230078	-	7.00	112.51	5.1e+03	282	BB230078	BB230078	RIKEN	full	
gb_estt2:BM010299	+	7.00	113.30	4.6e+03	256	BM010299	603631178F1	NIH_MGC_41		gb_gss:BB8042	-	7.00	112.51	5.1e+03	282	BB8042	F23M187B	IGF	Arbido	
gb_estt1:AV077072	-	7.00	113.23	4.6e+03	258	AV077072	AV077072	Mus	muscul	gb_estt1:AW038582	-	7.00	112.49	5.1e+03	283	AW038582	SNOWCMCA36601SK	On		
gb_estt2:BI681709	-	7.00	113.23	4.6e+03	258	BI681709	461166	MARC	1Bov	Bos	gb_gss:BB166916	-	7.00	112.49	5.1e+03	283	BB166916	SLK_000344	Arbido	
gb_estt1:AV055540	-	7.00	113.20	4.7e+03	259	AV055540	AV055540	Mus	muscul	gb_estt1:AA355960	-	7.00	112.46	5.1e+03	284	AA355960	EST654438	Jurkat	T-C	
gb_estt1:AV129179	-	7.00	113.20	4.7e+03	259	AV129179	AV129179	Mus	muscul	gb_estt1:BB154575	-	7.00	112.46	5.1e+03	284	BB154575	BB154575	Richt	hemt	
gb_estt2:FI14062	-	7.00	113.20	4.7e+03	259	FI14062	ATTS4984	Versailles-VB	A	gb_estt2:BF889065	+	7.00	112.46	5.1e+03	284	BF889065	RC4-RN0130-281100-0			
gb_estt1:AA207235	-	7.00	113.17	4.7e+03	260	AA207235	zq373g03.r1	Stratagene		gb_estt1:AV035655	+	7.00	112.43	5.1e+03	285	AV035655	AV035655	Mus	muscul	
gb_estt1:AV128125	-	7.00	113.17	4.7e+03	260	AV128125	AV128125	Mus	muscul	gb_estt1:AV209994	-	7.00	112.43	5.1e+03	285	AV209994	AV209994	RIKEN	full	
gb_estt1:AV292569	-	7.00	113.17	4.7e+03	260	AV292569	AV292569	RIKEN	full-16	gb_estt1:BB003919	-	7.00	112.43	5.1e+03	285	BB003919	BB003919	RIKEN	full	
gb_estt1:BB427321	-	7.00	113.17	4.7e+03	260	BB427321	BB427321	RIKEN	full-16	gb_estt1:BB009866	-	7.00	112.43	5.1e+03	285	BB009866	BB009866	RIKEN	full	
gb_estt2:BG946306	+	7.00	113.17	4.7e+03	260	BG946306	PM2-KR0007-120101-002-			gb_estt1:BB195185	-	7.00	112.43	5.1e+03	285	BB195185	BB195185	RIKEN	full	
gb_estt1:AV038524	-	7.00	113.14	4.7e+03	261	AV038524	AV038524	Mus	muscul	gb_estt1:BB287796	-	7.00	112.43	5.1e+03	285	BB287796	BB287796	RIKEN	full	
gb_estt2:BG992446	+	7.00	113.14	4.7e+03	261	BG992446	MR3-HR0999-020201-002-			gb_estt1:AV106104	-	7.00	112.37	5.2e+03	287	AV106104	AV106104	Mus	muscul	
gb_estt1:AV0397728	-	7.00	113.14	4.7e+03	261	AV0397728	HS_3040_B1	B08_MR	CIT	gb_estt1:BB233066	-	7.00	112.37	5.2e+03	287	BB233066	BB233066	RIKEN	full	
gb_estt1:AV131730	-	7.00	113.11	4.7e+03	262	AV131730	AV131730	Mus	muscul	gb_estt1:BB262307	-	7.00	112.37	5.2e+03	287	BB262307	BB262307	RIKEN	full	
gb_estt1:AV293845	+	7.00	113.11	4.7e+03	262	AV293845	AV293845	RIKEN	full-16	gb_estt1:BI069198	-	7.00	112.37	5.2e+03	287	BI069198	VI58E07.s1	Knowles		
gb_estt1:AA407712	-	7.00	113.11	4.7e+03	262	AA407712	ESTN0966	Mouse	7.5	dpd	gb_estt2:BI069294	-	7.00	112.37	5.2e+03	287	BI069294	CO34P21U	Populus	st
gb_gss:AA286E100	-	7.00	113.11	4.7e+03	262	AA286E100	T. brice	sheared	gend	gb_estt2:BM115451	+	7.00	112.37	5.2e+03	287	BM115451	L0820C01-3	NIA	Mous	
gb_estt1:AV016067	-	7.00	113.08	4.7e+03	263	AV016067	AV016067	Mus	muscul	gb_estt2:BF469294	+	7.00	112.37	5.2e+03	287	BF469294	UT-M-BH3-aug-d-12-0			
gb_estt1:AV120233	-	7.00	113.08	4.7e+03	263	AV120233	AV120233	Mus	muscul	gb_estt1:AV037189	-	7.00	112.34	5.2e+03	288	AV037189	AV037189	Mus	muscul	
gb_estt1:BB279453	-	7.00	113.08	4.7e+03	263	BB279453	BB279453	RIKEN	full-16	gb_estt1:BB292094	-	7.00	112.34	5.2e+03	288	BB292094	BB292094	RIKEN	full	
gb_estt1:BB384755	-	7.00	113.08	4.7e+03	263	BB384755	BB384755	RIKEN	full-16	gb_estt1:BB301775	-	7.00	112.34	5.2e+03	288	BB301775	BB301775	RIKEN	full	
gb_estt2:BM098347	-	7.00	113.08	4.7e+03	263	BM098347	EBem08.S0002.J13	R	IGH	gb_estt1:AV125360	-	7.00	112.32	5.2e+03	289	AV125360	AV125360	Mus	muscul	
gb_estt2:BF402323	-	7.00	113.08	4.7e+03	263	BF402323	UT-R-CA0-bhy-h-04-0	U1		gb_estt1:AV129353	-	7.00	112.32	5.2e+03	289	AV129353	AV129353	Mus	muscul	
gb_estt1:AV240748	-	7.00	113.05	4.7e+03	264	AV240748	AV240748	RIKEN	full-16	gb_estt1:AV149430	-	7.00	112.32	5.2e+03	289	AV149430	AV149430	Mus	muscul	
gb_estt1:AA260072	-	7.00	113.05	4.7e+03	264	AA260072	va38g11.r1	Soares	mous	gb_estt1:AV156270	-	7.00	112.32	5.2e+03	289	AV156270	AV156270	Mus	muscul	
gb_estt1:BB369666	-	7.00	113.05	4.7e+03	264	BB369666	BB369666	RIKEN	full-16	gb_estt1:BB349031	-	7.00	112.32	5.2e+03	289	BB349031	BB349031	RIKEN	full	
gb_estt2:HS201539	-	7.00	113.05	4.7e+03	264	HS201539	HS201539	Hom	SP16	gb_estt1:AV000779	-	7.00	112.29	5.2e+03	290	AV000779	AV000779	Mus	muscul	
gb_estt1:AV029155	-	7.00	113.05	4.7e+03	264	AV029155	RPCT-24-164P14.TJ	RPCT		gb_estt1:AV149267	-	7.00	112.29	5.2e+03	290	AV149267	AV149267	Mus	muscul	
gb_gss:AA2912545	+	7.00	113.05	4.7e+03	264	AA2912545	RPCT-24-164P14.TJ	RPCT		gb_estt1:AA083873	-	7.00	112.29	5.2e+03	290	AA083873	3316	MARC	1Bov	Bos
gb_estt1:AA108778	-	7.00	113.02	4.8e+03	265	AA108778	m35h07.r1	Barstead	MF	gb_estt1:BB008607	-	7.00	112.29	5.2e+03	290	BB008607	BB00607	RIKEN	full	
gb_estt2:BF847558	-	7.00	113.02	4.8e+03	265	BF847558	CM3-ENM048	311000-441-		gb_estt1:BB0232108	-	7.00	112.29	5.2e+03	290	BB0232108	BB232108	RIKEN	full	
gb_estt1:AV313255	-	7.00	112.99	4.8e+03	266	AV313255	AV313255	RIKEN	full-16	gb_estt1:BB232108	-	7.00	112.29	5.2e+03	290	BB232108	BB232108	RIKEN	full	
gb_estt1:AV362527	-	7.00	112.99	4.8e+03	266	AV362527	AV362527	RIKEN	full-16	gb_estt1:BB232108	-	7.00	112.29	5.2e+03	290	BB232108	BB232108	RIKEN	full	
gb_estt2:BF290155	-	7.00	112.99	4.8e+03	266	BF290155	EST9454746	Rat	Gene	Ind	gb_estt2:J2826781	-	7.00	112.29	5.2e+03	290	J2826781	EST9454746	Human	Embryo
gb_estt1:AV042913	-	7.00	112.93	4.8e+03	268	AV042913	AV042913	Mus	muscul	gb_gss:BBH31813	-	7.00	112.29	5.2e+03	290	BBH31813	CH230-40M14.TJ	CHOR		
gb_estt2:BB925248	-	7.00	112.90	4.8e+03	269	BB925248	CM4-ANM081	210800-274-		gb_estt1:AV037202	-	7.00	112.26	5.3e+03	291	AV037202	AV037202	Mus	muscul	
gb_estt1:HS32104	-	7.00	112.90	4.8e+03	269	HS32104	EST106901	Rat	PC-12	cell	gb_estt1:AV289926	-	7.00	112.26	5.3e+03	291	AV289926	AV299926	RIKEN	full
gb_gss:AA077876	-	7.00	112.90	4.8e+03	269	AA077876	CIT-HSP-236662.TF	CIT		gb_estt1:BB388463	-	7.00	112.26	5.3e+03	291	BB388463	BB388463	RIKEN	full	
gb_estt1:AA1428009	-	7.00	112.87	4.9e+03	270	AA1428009	mK09a06.x1	Soares	mous	gb_estt1:BB355521	-	7.00	112.26	5.3e+03	291	BB355521	CH230-69F3.TV	CHORI		
gb_estt1:AA1659816	-	7.00	112.87	4.9e+03	270	AA1659816	tU03c09.x1	NCI	CGAP	PI	gb_gss:BBH35521	-	7.00	112.26	5.3e+03	291	BBH35521	CH230-69F3.TV	CHORI	
gb_estt1:AV127766	-	7.00	112.87	4.9e+03	270	AV127766	AV127766	Mus	muscul	gb_estt1:AV0899106	-	7.00	112.23	5.3e+03	292	AV0899106	AV089106	Mus	muscul	
gb_gss:BB33406	+	7.00	112.87	4.9e+03	270	BB33406	HS-1017-B1-D04-MF	abi	CT	gb_estt1:AV133444	-	7.00	112.23	5.3e+03	292	AV133444	AV133444	Mus	muscul	
gb_gss:BBH333073	-	7.00	112.87	4.9e+03	270	BBH333073	CH230-202A18.TJ	CHORI		gb_estt1:BB478372	-	7.00	112.23	5.3e+03	292	BB478372	BB478372	RIKEN	full	
gb_estt2:BB477362	-	7.00	112.84	4.9e+03	271	BB477362	us51d08.y1	Perkins	LRI	gb_estt1:AV208193	-	7.00	112.20	5.3e+03	293	AV208193	AV208193	RIKEN	full	
gb_estt1:AA303636	-	7.00	112.81	4.9e+03	272	AA303636	EST16609	Aorta	endoth	gb_estt1:BB004755	-	7.00	112.20	5.3e+03	293	BB004755	BB004755	RIKEN	full	
gb_estt1:AAW945614	-	7.00	112.81	4.9e+03	272	AAW945614	RC2-ENM027-290400-011-			gb_estt1:BB068179	-	7.00	112.20	5.3e+03	293	BB068179	BB068179	RIKEN	full	
gb_gss:AA246929	-	7.00	112.75	4.9e+03	273	AA246929	IM0243M02R	Mouse	10kb	gb_estt1:AA890868										

gb_gss: A2364844	7.00	112.12	5.3e+03	296	A2364844 IM011113f Mouse 10kb	gb_est2: C35334	7.00	112.01	5.4e+03	300	C35334 C35334 Yuj1 Kohara un
gb_est1: AV034688	7.00	112.09	5.4e+03	297	AV034688 AV034688 Mus musculus	gb_est2: C35340	7.00	112.01	5.4e+03	300	C35340 C35340 Yuj1 Kohara un
gb_est1: BB339322	7.00	112.09	5.4e+03	297	BB339322 BB339322 Riken full-1f	gb_est2: C36101	7.00	112.01	5.4e+03	300	C36101 C36101 Yuj1 Kohara un
gb_est1: BB555278	7.00	112.09	5.4e+03	297	BB555278 BB555278 Riken full-1f	gb_est2: C36780	7.00	112.01	5.4e+03	300	C36780 C36780 Yuj1 Kohara un
gb_est2: T01098	7.00	112.09	5.4e+03	297	T01098 WEST01819 Early embryo,	gb_est2: C36806	7.00	112.01	5.4e+03	300	C36806 C36806 Yuj1 Kohara un
gb_est1: AV104510	7.00	112.07	5.4e+03	298	AV104510 AV104510 Mus musculus	gb_est2: C36846	7.00	112.01	5.4e+03	300	C36846 C36846 Yuj1 Kohara un
gb_est1: BB178368	7.00	112.07	5.4e+03	298	BB178368 AV148368 Mus musculus	gb_est2: C36957	7.00	112.01	5.4e+03	300	C36957 C36957 Yuj1 Kohara un
gb_est1: BB177213	7.00	112.07	5.4e+03	298	BB177213 BB177213 Riken full-1f	gb_est2: C37479	7.00	112.01	5.4e+03	300	C37479 C37479 Yuj1 Kohara un
gb_est2: BF661535	7.00	112.07	5.4e+03	298	BF661535 d682f03r y1 Wellcome CR	gb_est2: C37774	7.00	112.01	5.4e+03	300	C37774 C37774 Yuj1 Kohara un
gb_est1: AV040324	7.00	112.04	5.4e+03	299	AV040324 AV040324 Mus musculus	gb_est2: C38064	7.00	112.01	5.4e+03	300	C38064 C38064 Yuj1 Kohara un
gb_est1: BB074970	7.00	112.04	5.4e+03	299	BB074970 BB074970 Riken full-1f	gb_est2: C38333	7.00	112.01	5.4e+03	300	C38333 C38333 Yuj1 Kohara un
gb_est1: BB217267	7.00	112.04	5.4e+03	299	BB217267 BB217267 Riken full-1f	gb_est2: C38330	7.00	112.01	5.4e+03	300	C38330 C38330 Yuj1 Kohara un
gb_est1: BB364408	7.00	112.04	5.4e+03	299	BB364408 BB364408 Riken full-1f	gb_est2: C38433	7.00	112.01	5.4e+03	300	C38433 C38433 Yuj1 Kohara un
gb_est1: BB554257	7.00	112.04	5.4e+03	299	BB554257 BB554257 Riken full-1f	gb_est2: C39151	7.00	112.01	5.4e+03	300	C39151 C39151 Yuj1 Kohara un
gb_est1: AV098764	7.00	112.01	5.4e+03	300	AV098764 AV098764 Sugeno Homo S	gb_gss: A244428	7.00	112.01	5.4e+03	300	A244428 IM0239p10r Mouse 10
gb_est1: AV113694	7.00	112.01	5.4e+03	300	AV113694 AV113694 unpublished c	gb_est2: H33584	7.00	111.99	5.4e+03	301	H33584 EST109725 Rat PC-12 c
gb_est1: AV115487	7.00	112.01	5.4e+03	300	AV115487 AV115487 unpublished c	gb_est1: BB220759	7.00	111.96	5.5e+03	302	BB220759 BB220759 Riken full1
gb_est1: AV096384	7.00	112.01	5.4e+03	300	AV096384 AV096384 Mus musculus	gb_est1: BB223027	7.00	111.96	5.5e+03	302	BB223027 BB223027 Riken full1
gb_est1: AV151945	7.00	112.01	5.4e+03	300	AV151945 AV151945 Mus musculus	gb_est1: BB516815	7.00	111.96	5.5e+03	302	BB516815 BB516815 Riken full1
gb_est1: AV175870	7.00	112.01	5.4e+03	300	AV175870 AV175870 Yuj1 Kohara u	gb_est1: AA221917	7.00	111.93	5.5e+03	303	AA221917 mw16h01_r1 Soares m
gb_est1: AV176191	7.00	112.01	5.4e+03	300	AV176191 AV176191 Yuj1 Kohara u	gb_est1: BB120926	7.00	111.93	5.5e+03	303	BB120926 BB120926 Riken full1
gb_est1: AV176665	7.00	112.01	5.4e+03	300	AV176665 AV176665 Yuj1 Kohara u	gb_est1: BB477842	7.00	111.93	5.5e+03	303	BB477842 BB477842 Riken full1
gb_est1: AV176836	7.00	112.01	5.4e+03	300	AV176836 AV176836 Yuj1 Kohara u	gb_est1: BB809793	7.00	111.93	5.5e+03	303	BB809793 BB809793 Riken full1
gb_est1: AV176962	7.00	112.01	5.4e+03	300	AV176962 AV176962 Yuj1 Kohara u	gb_gss: A2465428	7.00	111.93	5.5e+03	304	A2465428 IM0275b18f Mouse 10
gb_est1: AV177696	7.00	112.01	5.4e+03	300	AV177696 AV177696 Yuj1 Kohara u	gb_est1: AV055296	7.00	111.91	5.5e+03	304	AV055296 AV055296 Mus muscul
gb_est1: AV178568	7.00	112.01	5.4e+03	300	AV178568 AV178568 Yuj1 Kohara u	gb_est1: AV082633	7.00	111.91	5.5e+03	304	AV082633 AV082633 Mus muscul
gb_est1: AV179041	7.00	112.01	5.4e+03	300	AV179041 AV179041 Yuj1 Kohara u	gb_est1: AV121146	7.00	111.91	5.5e+03	304	AV121146 AV121146 Mus muscul
gb_est1: AV179289	7.00	112.01	5.4e+03	300	AV179289 AV179289 Yuj1 Kohara u	gb_est2: D62486	7.00	111.91	5.5e+03	304	D62486 HMC29D07B Clontech h
gb_est1: AV179316	7.00	112.01	5.4e+03	300	AV179316 AV179316 Yuj1 Kohara u	gb_est2: Z44674	7.00	111.91	5.5e+03	304	Z44674 HSC265031 normalized
gb_est1: AV179951	7.00	112.01	5.4e+03	300	AV179951 AV179951 Yuj1 Kohara u	gb_gss: A0907967	7.00	111.91	5.5e+03	304	A0907967 GSTR05334 Trypanos
gb_est1: AV180168	7.00	112.01	5.4e+03	300	AV180168 AV180168 Yuj1 Kohara u	gb_gss: BH294987	7.00	111.91	5.5e+03	304	BH294987 CH230-44L16 Tj CHOR
gb_est1: AV180641	7.00	112.01	5.4e+03	300	AV180641 AV180641 Yuj1 Kohara u	gb_est1: A465143	7.00	111.88	5.5e+03	305	A465143 mw74h09_r1 Soares m
gb_est1: AV180813	7.00	112.01	5.4e+03	300	AV180813 AV180813 Yuj1 Kohara u	gb_est1: A561735	7.00	111.88	5.5e+03	305	A561735 v662h03_r1 Soares t
gb_est1: AV180845	7.00	112.01	5.4e+03	300	AV180845 AV180845 Yuj1 Kohara u	gb_est1: AV060593	7.00	111.88	5.5e+03	305	AV060593 AV060593 Mus muscul
gb_est1: AV180962	7.00	112.01	5.4e+03	300	AV180962 AV180962 Yuj1 Kohara u	gb_est1: AV099888	7.00	111.88	5.5e+03	305	AV099888 AV099888 Mus muscul
gb_est1: AV181070	7.00	112.01	5.4e+03	300	AV181070 AV181070 Yuj1 Kohara u	gb_est2: BE451733	7.00	111.88	5.5e+03	305	BE451733 NXC1_001.C08_F NCBI
gb_est1: AV181082	7.00	112.01	5.4e+03	300	AV181082 AV181082 Yuj1 Kohara u	gb_est2: BE928966	7.00	111.88	5.5e+03	305	BE928966 PM3-GN0012-280800-0
gb_est1: AV181124	7.00	112.01	5.4e+03	300	AV181124 AV181124 Yuj1 Kohara u	gb_est2: BE928999	7.00	111.85	5.5e+03	305	BE928999 PM3-GN0012-290800-0
gb_est1: AV181807	7.00	112.01	5.4e+03	300	AV181807 AV181807 Yuj1 Kohara u	gb_est1: AV038497	7.00	111.85	5.5e+03	306	AV038497 AV038497 Dictyostel
gb_est1: AV182339	7.00	112.01	5.4e+03	300	AV182339 AV182339 Yuj1 Kohara u	gb_est1: AM540739	7.00	111.83	5.6e+03	307	AM540739 C0136P03-3 NTA Mous
gb_est1: AV182749	7.00	112.01	5.4e+03	300	AV182749 AV182749 Yuj1 Kohara u	gb_est1: AV027493	7.00	111.80	5.6e+03	308	AV027493 AV027493 Mus muscul
gb_est1: AV183431	7.00	112.01	5.4e+03	300	AV183431 AV183431 Yuj1 Kohara u	gb_est1: AV166744	7.00	111.80	5.6e+03	308	AV166744 AV166744 Mus muscul
gb_est1: AV183656	7.00	112.01	5.4e+03	300	AV183656 AV183656 Yuj1 Kohara u	gb_gss: BH359127	7.00	111.80	5.6e+03	308	BH359127 CH230-GEF22 TV CHORI
gb_est1: AV183691	7.00	112.01	5.4e+03	300	AV183691 AV183691 Yuj1 Kohara u	gb_gss: TH050709	7.00	111.80	5.6e+03	308	TH050709 PM500709 Riken full1
gb_est1: AV184110	7.00	112.01	5.4e+03	300	AV184110 AV184110 Yuj1 Kohara u	gb_gss: AV112515	7.00	111.72	5.6e+03	311	AV112515 AV112515 Mus muscul
gb_est1: AV184698	7.00	112.01	5.4e+03	300	AV184698 AV184698 Yuj1 Kohara u	gb_est1: BE165241	7.00	111.72	5.6e+03	311	BE165241 PM1-H70521-050500-0
gb_est1: AV184870	7.00	112.01	5.4e+03	300	AV184870 AV184870 Yuj1 Kohara u	gb_est1: AM833376	7.00	111.77	5.6e+03	309	AM833376 G039G03_r1 NCI_CGAP
gb_est1: AV185561	7.00	112.01	5.4e+03	300	AV185561 AV185561 Yuj1 Kohara u	gb_est1: BB203028	7.00	111.77	5.6e+03	309	BB203028 BB203028 Riken full1
gb_est1: AV880761	7.00	112.01	5.4e+03	300	AV880761 AV880761 Nori Satoch un	gb_est1: BB607366	7.00	111.75	5.6e+03	309	BB607366 BB607366 Riken full1
gb_est1: BB241194	7.00	112.01	5.4e+03	300	BB241194 BB241194 Riken full-1f	gb_est1: AV112515	7.00	111.72	5.6e+03	311	AV112515 AV112515 Mus muscul
gb_est1: C07987	7.00	112.01	5.4e+03	300	C07987 C07987 Yuj1 Kohara unpub	gb_est2: BE646285	7.00	111.70	5.6e+03	312	BE646285 T68hB06_r1 NCI_CGAP
gb_est2: C296424	7.00	112.01	5.4e+03	300	C296424 C296424 Yuj1 Kohara unpub	gb_gss: A2070670	7.00	111.70	5.6e+03	312	A2070670 RPECI-23-421C16_Tj CH
gb_est2: C29665	7.00	112.01	5.4e+03	300	C29665 C29665 Yuj1 Kohara unpub	gb_gss: BH35376	7.00	111.59	5.7e+03	316	BH35376 CH230-173624_TV CHO
gb_est2: C29665	7.00	112.01	5.4e+03	300	C29665 C29665 Yuj1 Kohara unpub	gb_gss: AA385856	7.00	111.57	5.7e+03	317	AA385856 E9996633 Thyroid Ho
gb_est2: C29759	7.00	112.01	5.4e+03	300	C29759 C29759 Yuj1 Kohara unpub	gb_gss: AV133072	7.00	111.54	5.8e+03	318	AV133072 ZM0001007r Mouse 10
gb_est2: C29857	7.00	112.01	5.4e+03	300	C29857 C29857 Yuj1 Kohara unpub	gb_est1: AV130072	7.00	111.54	5.8e+03	318	AV130072 AV130072 Mus muscul
gb_est2: C29959	7.00	112.01	5.4e+03	300	C29959 C29959 Yuj1 Kohara unpub	gb_est1: AV208145	7.00	111.54	5.8e+03	318	AV208145 AV208145 Riken full1
gb_est2: C29966	7.00	112.01	5.4e+03	300	C29966 C29966 Yuj1 Kohara unpub	gb_est1: AA288462	7.00	111.54	5.8e+03	318	AA288462 v16f611_r1 Soares m
gb_est2: C30318	7.00	112.01	5.4e+03	300	C30318 C30318 Yuj1 Kohara unpub	gb_gss: BH35376	7.00	111.52	5.8e+03	318	BH35376 CH230-173624_TV CHO
gb_est2: C30410	7.00	112.01	5.4e+03	300	C30410 C30410 Yuj1 Kohara unpub	gb_gss: A4141944	7.00	111.52	5.8e+03	319	A4141944 s983a07_r1 Gm-cl004
gb_est2: C30517	7.00	112.01	5.4e+03	300	C30517 C30517 Yuj1 Kohara unpub	gb_est1: A1723172	7.00	111.52	5.8e+03	319	A1723172 1071 Ptf6f2 Plaus t
gb_est2: C30651	7.00	112.01	5.4e+03	300	C30651 C30651 Yuj1 Kohara unpub	gb_est2: BE986015	7.00	111.52	5.8e+03	319	BE986015 CML-D70048-170101-7
gb_est2: C30783	7.00	112.01	5.4e+03	300	C30783 C30783 Yuj1 Kohara unpub	gb_est2: BE944188	7.00	111.49	5.8e+03	320	BE944188 ujf-BH3-4wn-h-05-0
gb_est2: C30806	7.00	112.01	5.4e+03	300	C30806 C30806 Yuj1 Kohara unpub	gb_est1: AA064282	7.00	111.49	5.8e+03	320	AA064282 m167-f06_r1 Soares m
gb_est2: C31209	7.00	112.01	5.4e+03	300	C31209 C31209 Yuj1 Kohara unpub	gb_est1: A1034580	7.00	111.49	5.8e+03	320	A1034580 Lmk339p10/252c Lets
gb_est2: C32518	7.00	112.01	5.4e+03	300	C32518 C32518 Yuj1 Kohara unpub	gb_est1: BB136940	7.00	111.49	5.8e+03	320	BB136940 BB136940 Riken full1
gb_est2: C32966	7.00	112.01	5.4e+03	300	C32966 C32966 Yuj1 Kohara unpub	gb_est1: BB140572	7.00	111.49	5.8e+03	320	BB140572 BB140572 Riken full1
gb_est2: C32994	7.00	112.01	5.4e+03	300	C32994 C32994 Yuj1 Kohara unpub						
gb_est2: C33492	7.00	112.01	5.4e+03	300	C33492 C33492 Yuj1 Kohara unpub						
gb_est2: C34761	7.00	112.01	5.4e+03	300	C34761 C34761 Yuj1 Kohara unpub						
gb_est2: C34875	7.00	112.01	5.4e+03	300	C34875 C34875 Yuj1 Kohara unpub						
gb_est2: C35145	7.00	112.01	5.4e+03	300	C35145 C35145 Yuj1 Kohara unpub						
gb_est2: C35185	7.00	112.01	5.4e+03	300	C35185 C35185 Yuj1 Kohara unpub						

gb_est2:BF802920	-	7.00	111.49	5.8e+03	320	BF802920	IL5-C10149-011100-222-	gb_est1:AV209743	+	7.00	110.90	6.2e+03	344	AV209743	AN209743	RIKEN	full
gb_est2:BI133335	+	7.00	111.49	5.8e+03	320	BI133335	UT-M-BH3-brv-g-06-0-01	gb_est2:BF339943	-	7.00	110.90	6.2e+03	344	BF339943	UT-R-CAL-Bjb-h-17-0		
gb_est1:AA736935	+	7.00	111.46	5.8e+03	321	AA736935	oa48c03.s1 NC1 CGAP GC	gb_est1:AV011452	-	7.00	110.86	6.3e+03	346	AV011452	AV011452	Mus	muscul
gb_est1:AAV039711	+	7.00	111.46	5.8e+03	321	AAV039711	476 PIPIFG plus tede	gb_est2:237589	-	7.00	110.86	6.3e+03	346	237589	ATTS4054	Strasbourg-A	
gb_est1:AV090316	+	7.00	111.46	5.8e+03	321	AV090316	AV090316	gb_est2:BE748459	-	7.00	110.86	6.3e+03	346	BE748459	601571953PI	NIH-MGC	
gb_est2:DI71307	+	7.00	111.46	5.8e+03	321	DI71307	CELMK067CYR	gb_gss:AE514891	-	7.00	110.86	6.3e+03	346	AE514891	IM0361E2AR	Mouse	10
gb_est1:AA212313	+	7.00	111.44	5.8e+03	322	AA212313	ma77d04.r1 Soares	gb_est1:BI202100	-	7.00	110.83	6.3e+03	347	BI202100	q152a04.r1	NC1-CGAP	
gb_est1:AAW012518	+	7.00	111.44	5.8e+03	322	AAW012518	IL5-DU0067-030400-053-	gb_est2:BE638842	-	7.00	110.83	6.3e+03	347	BE638842	RPCT-23-209113	TV-R	
gb_est1:BB140573	-	7.00	111.44	5.8e+03	322	BB140573	BB140573	gb_gss:AB555442	+	7.00	110.83	6.3e+03	347	AB555442	RPCT-11-3111	TV-R	
gb_est2:BM139502	-	7.00	111.44	5.8e+03	322	BM139502	k165f04.y1 Ascatis	gb_gss:BA48779	+	7.00	110.81	6.3e+03	348	BA48779	me77f02.r1	Soares	mon
gb_est2:HB98464	-	7.00	111.44	5.8e+03	322	HB98464	yw04e05.r1 Soares	gb_est2:W82840	-	7.00	110.81	6.3e+03	348	W82840	me77f02.r1	Soares	mon
gb_gss:AZ706922	-	7.00	111.44	5.8e+03	322	AZ706922	RPCT-23-23682	gb_est2:BF5542284	+	7.00	110.81	6.3e+03	348	BF5542284	UT-R-C2-rs-C-06-0-0		
gb_gss:BM363863	-	7.00	111.44	5.8e+03	322	BM363863	CH230-119012	gb_est1:AA441046	+	7.00	110.79	6.3e+03	349	AA441046	v166g03.r1	Soares	pla
gb_est1:AV161696	-	7.00	111.41	5.9e+03	323	AV161696	AV161696	gb_est2:K233372	+	7.00	110.79	6.3e+03	349	K233372	yh32901.r1	Soares	pla
gb_est1:AA1900133	-	7.00	111.39	5.9e+03	324	AA1900133	w29e08.x1 NC1 CGAP GC	gb_gss:AZ383245	+	7.00	110.79	6.3e+03	349	AZ383245	IM014910AR	Mouse	10
gb_est2:DI71465	+	7.00	111.36	5.9e+03	324	DI71465	CELMK0755LR	gb_gss:AO247626	-	7.00	110.79	6.3e+03	349	AO247626	HS-2055	B2_F09_T7	C
gb_est2:DA24230	+	7.00	111.36	5.9e+03	324	DA24230	BB224230	gb_gss:AO294832	-	7.00	110.79	6.3e+03	349	AO294832	HS-2183	B1_H12_MF	C
gb_est2:BG982818	+	7.00	111.34	5.9e+03	326	BG982818	PMO-CN0150-010301-001	gb_gss:AO297085	-	7.00	110.79	6.3e+03	349	AO297085	HS-3034	B2_C10_MF	C
gb_est2:BE517111	+	7.00	111.34	5.9e+03	326	BE517111	WHE0624_C07_F142A	gb_est2:BG053004	+	7.00	110.76	6.4e+03	350	BG053004	H3003A03-3	NIA	Mous
gb_gss:AZ617949	+	7.00	111.34	5.9e+03	326	AZ617949	IM044024F	gb_est1:BA5288942	+	7.00	110.74	6.4e+03	351	BA5288942	AV528892	Arbidiops	1
gb_est1:AA408233	+	7.00	111.31	5.9e+03	327	AA408233	EST02735	gb_est1:BA528892	+	7.00	110.74	6.4e+03	351	BA528892	CM3-H70192-071099-0		
gb_gss:AO918294	+	7.00	111.31	5.9e+03	327	AO918294	RPCT-23-287R	gb_est1:BE145088	-	7.00	110.74	6.4e+03	351	BE145088	CM3-H70192-071099-0		
gb_gss:BM314949	+	7.00	111.29	5.9e+03	327	BM314949	CH230-119G16	gb_gss:AO933894	-	7.00	110.74	6.4e+03	351	AO933894	RPCT-23-324H15	TJ R	
gb_est1:BB640646	+	7.00	111.29	5.9e+03	328	BB640646	BB640646	gb_gss:AO986534	-	7.00	110.74	6.4e+03	351	AO986534	HS-3041	B1_A03_LMR	C
gb_est1:BB809585	+	7.00	111.29	5.9e+03	328	BB809585	BB809585	gb_gss:AO301772	-	7.00	110.74	6.4e+03	351	AO301772	HS-3041	B1_A03_LMR	C
gb_est2:BI0235356	+	7.00	111.29	5.9e+03	328	BI0235356	CM4-MW0247-190101-815-	gb_est2:BI338648	-	7.00	110.72	6.4e+03	352	BI338648	MI-P-AY1-npv-c-09-0		
gb_est2:HB87742	+	7.00	111.29	5.9e+03	328	HB87742	y574f06.s1 Soares	gb_est1:BI824517	-	7.00	110.70	6.4e+03	353	BI824517	lx71c08.x1	NC1-CGAP	
gb_est2:BE522777	+	7.00	111.27	5.9e+03	329	BE522777	M28G658M	gb_est1:AU070993	-	7.00	110.70	6.4e+03	353	AU070993	AU070993	Stracena	
gb_est2:DI72554	+	7.00	111.27	6.0e+03	329	DI72554	CELMK1046R	gb_est2:394714	-	7.00	110.67	6.4e+03	353	394714	y635610	r1	Stracena
gb_est1:AA106152	+	7.00	111.24	6.0e+03	330	AA106152	hb02a12_p1	gb_est1:MW346576	-	7.00	110.67	6.4e+03	354	MW346576	28441	MARC	PIFG
gb_est1:AV547237	+	7.00	111.24	6.0e+03	330	AV547237	Arbidiops	gb_gss:BA324106	-	7.00	110.67	6.4e+03	354	BA324106	CH230-15624	TJ CHOR	
gb_est2:BF683824	+	7.00	111.24	6.0e+03	330	BF683824	QV1-MW0170-131100-459-	gb_est1:AV014209	+	7.00	110.65	6.5e+03	355	AV014209	OVI-BN0001-30A1199-0		
gb_gss:AM074727	+	7.00	111.24	6.0e+03	330	AM074727	CIT-HSP-23514A	gb_est1:AM747893	-	7.00	110.65	6.5e+03	355	AM747893	OVI-BN0001-30A1199-0		
gb_gss:BM171755	+	7.00	111.24	6.0e+03	330	BM171755	SALK_004749	gb_gss:AO650478	-	7.00	110.65	6.5e+03	355	AO650478	Sheared DNA-20J5	TR	
gb_est1:AA145738	+	7.00	111.17	6.0e+03	332	AA145738	q43d09.y1 Moss	gb_est1:AA247820	-	7.00	110.63	6.5e+03	356	AA247820	HF-00711	seq.f	Homan
gb_est1:BB144612	+	7.00	111.17	6.0e+03	333	BB144612	BB144612	gb_est1:AA247820	-	7.00	110.63	6.5e+03	356	AA247820	HF-00711	seq.f	Homan
gb_est1:BB145025	+	7.00	111.17	6.0e+03	333	BB145025	BB145025	gb_est1:AA247820	-	7.00	110.63	6.5e+03	356	AA247820	HF-00711	seq.f	Homan
gb_est1:BB254983	+	7.00	111.17	6.0e+03	333	BB254983	BB254983	gb_est1:AA247820	-	7.00	110.63	6.5e+03	356	AA247820	HF-00711	seq.f	Homan
gb_gss:AO522302	+	7.00	111.17	6.0e+03	333	AO522302	HS-5207	gb_est1:AA247820	-	7.00	110.63	6.5e+03	356	AA247820	HF-00711	seq.f	Homan
gb_est1:AAW422146	+	7.00	111.14	6.1e+03	334	AAW422146	T210118e	gb_est1:AA247820	-	7.00	110.63	6.5e+03	356	AA247820	HF-00711	seq.f	Homan
gb_est2:BB882372	-	7.00	111.14	6.1e+03	334	BB882372	CM1-EM0192-051200-621-	gb_est1:AAW47543	-	7.00	110.60	6.5e+03	357	AAW47543	UT-M-BH1-ama-f-02-0		
gb_est1:BB456940	+	7.00	111.12	6.1e+03	335	BB456940	BB456940	gb_est1:AAW47543	-	7.00	110.60	6.5e+03	357	AAW47543	UT-M-BH1-ama-f-02-0		
gb_est2:BI595286	+	7.00	111.12	6.1e+03	335	BI595286	HC-L3-04B05	gb_gss:BA261117	-	7.00	110.60	6.5e+03	357	BA261117	CH230-170P10	TV CHO	
gb_est2:ID72180	+	7.00	111.09	6.1e+03	336	ID72180	CELMK01A2R	gb_est1:AV101954	+	7.00	110.58	6.5e+03	358	AV101954	AV101954	Mus	muscul
gb_est2:FB01157	+	7.00	111.09	6.1e+03	336	FB01157	HSBA6E112	gb_est1:BM413026	-	7.00	110.58	6.5e+03	358	BM413026	EST587353	tomato	br
gb_est2:BV19922	+	7.00	111.07	6.1e+03	337	BV19922	BB519922	gb_est1:AM797487	-	7.00	110.56	6.5e+03	359	AM797487	CM1-UW0039-030400-1		
gb_est2:DI71206	+	7.00	111.07	6.1e+03	337	DI71206	CELMK0352R	gb_est2:BE399705	-	7.00	110.56	6.5e+03	359	BE399705	WHE0041	G05P90703	
gb_est2:BE341500	+	7.00	111.07	6.1e+03	337	BE341500	EST393237	gb_est2:BE399705	-	7.00	110.56	6.5e+03	359	BE399705	WHE0041	G05P90703	
gb_est2:BF042962	+	7.00	111.07	6.1e+03	337	BF042962	BP250009B10D1	gb_est1:AI747967	-	7.00	110.54	6.6e+03	360	AI747967	IB49901.x1	Zebrafis	
gb_gss:AO884000	+	7.00	111.07	6.1e+03	337	AO884000	HS-5491	gb_gss:AO133056	-	7.00	110.54	6.6e+03	360	AO133056	AV193056	Yuj1	Kohar
gb_gss:BT9585	+	7.00	111.05	6.1e+03	337	BT9585	CIT-HSP-2045A9	gb_gss:BM2911795	-	7.00	110.51	6.6e+03	361	BM2911795	CH230-30E7	TV CHO	
gb_est1:AA865134	+	7.00	111.05	6.1e+03	338	AA865134	v683d06.r1	gb_est1:AA038044	-	7.00	110.51	6.6e+03	361	AA038044	IM80911	r1	Soares
gb_est1:AI047546	+	7.00	111.05	6.1e+03	338	AI047546	ub80b05.r1	gb_est1:AA111336	-	7.00	110.51	6.6e+03	361	AA111336	mp10a08.r1	liffe	Tec
gb_est1:AV0834756	+	7.00	111.05	6.1e+03	338	AV0834756	ub80b05.r1	gb_est1:AA111336	-	7.00	110.51	6.6e+03	361	AA111336	mp10a08.r1	liffe	Tec
gb_est1:BB223182	+	7.00	111.05	6.1e+03	338	BB223182	BB223182	gb_est2:BF372545	-	7.00	110.51	6.6e+03	361	BF372545	CM1-F70046	normaliz	
gb_est1:AA412430	+	7.00	111.05	6.1e+03	338	AA412430	zui10c09.r1	gb_gss:AO133056	-	7.00	110.51	6.6e+03	361	AO133056	AV193056	Yuj1	Kohar
gb_est2:W08934	+	7.00	111.05	6.1e+03	338	W08934	ma60f11.r1	gb_gss:AO133056	-	7.00	110.51	6.6e+03	361	AO133056	AV193056	Yuj1	Kohar
gb_gss:BM294582	-	7.00	111.05	6.1e+03	338	BM294582	CH230-173B3	gb_gss:AO243163	-	7.00	110.51	6.6e+03	361	AO243163	HS-2055	B2	DLI
gb_est1:AA920807	+	7.00	111.02	6.2e+03	339	AA920807	v68f07.r1	gb_gss:AA669494	-	7.00	110.49	6.6e+03	362	AA669494	v164b01.r1	Barstead	
gb_est2:DI71401	+	7.00	111.02	6.2e+03	339	DI71401	CELMK0742R	gb_est1:BE111051	+	7.00	110.49	6.6e+03	362	BE111051	UT-R-BJ1-auz-b-01-0		
gb_gss:AZ234544	+	7.00	111.00	6.2e+03	340	AZ234544	ATTS2859	gb_est2:BG577442	-	7.00	110.47	6.6e+03	363	BG577442	IM6	MRTZ	ADIS
gb_gss:AZ602747	+	7.00	111.00	6.2e+03	340	AZ602747	ATTS2859	gb_est1:WM113990	-	7.00	110.47	6.6e+03	363	WM113990	MC6540	mouse	liver
gb_est1:AAU85896	+	7.00	110.98	6.2e+03	341	AAU85896	AV085896	gb_est1:AA1616249	-	7.00	110.45	6.6e+03	364	AA1616249	v61e003.x1	Knowles	
gb_est1:AAU008618	+	7.00	110.98	6.2e+03	341	AAU008618	AV008618	gb_est1:AA1616249	-	7.00	110.45	6.6e+03	364	AA1616249	v61e003.x1	Knowles	
gb_est1:BB166676	+	7.00	110.98	6.2e+03													

gb_gss:CNS00Y00	7.00	110.40	6.7e+03	366	AL095574	Arabidopsis thaliana g	gb_est2:BE373770	7.00	110.10	6.9e+03	380	BE373770	601226176P1	NCI_CGA	
gb_gss:AO356091	7.00	110.40	6.7e+03	366	AO356091	CitriBt-E1-252B21.TR C	gb_gss:AA873958	7.00	110.10	6.9e+03	380	AA873958	2M018719R	mouse 10	
gb_est1:AM109667	7.00	110.38	6.7e+03	367	AM109667	MT11299 mouse liver c	gb_gss:AA0230114	7.00	110.10	6.9e+03	380	AA0230114	RPCT11-750L1	TK RPC	
gb_est2:BE427911	7.00	110.38	6.7e+03	367	BE427911	MTD001-1.B11990615 ID	gb_est1:AA690533	7.00	110.08	6.9e+03	381	AA690533	VC44b11.r1	Bartstead	
gb_gss:BF677303	7.00	110.38	6.7e+03	367	BF677303	602087009P1 NIH.MGC.83	gb_est1:AI937107	7.00	110.08	6.9e+03	381	AI937107	wp73b10.x1	NCI_CGAP	
gb_est1:AA900033	7.00	110.36	6.7e+03	367	AA900033	UI-R-EO-dh-f-02-0-UI-S	gb_est1:AA660890	7.00	110.08	6.9e+03	381	AA660890	WS7-7.C05.y1.A002 W		
gb_est1:AA936165	7.00	110.36	6.7e+03	368	AA936165	UI-R-EO-dh-f-02-0-UI-S	gb_est1:AA464421	7.00	110.08	6.9e+03	381	AA464421	zx78907.x1	Soares o	
gb_est1:AA545517	7.00	110.36	6.7e+03	368	AA545517	18339 MARC lBOV Bos t	gb_gss:AA256214	7.00	110.08	6.9e+03	381	AA256214	518529 MARC 3BOV Bo		
gb_est1:AA557016	7.00	110.36	6.7e+03	368	AA557016	CO194009-3 N1A Mouse E	gb_gss:AA825203	7.00	110.08	6.9e+03	381	AA825203	RPCT-23-466P1	TYB R	
gb_est2:WA8287	7.00	110.36	6.7e+03	368	WA8287	VO18907.r1	Barstead mc	gb_gss:AA3082224	7.00	110.08	6.9e+03	381	AA3082224	2M0262019R	mouse 10
gb_est1:AI183783	7.00	110.34	6.7e+03	369	AI183783	qei18903.x1	Soares_fete	gb_est1:AI1182337	7.00	110.06	7.0e+03	382	AI1182337	qdt6e07.x1	Soares_t
gb_est2:BF731622	7.00	110.34	6.7e+03	369	BF731622	mab89007.y1	NCI_CGAP-S	gb_est1:AI1283300	7.00	110.06	7.0e+03	382	AI1283300	qv08e10.x1	NCI_CGAP
gb_est2:7006870	7.00	110.34	6.7e+03	369	7006870	EST01331	Early embryo,	gb_est2:AA021811	7.00	110.06	7.0e+03	382	AA021811	K121b09.y1	Ascaris t
gb_est2:7082167	7.00	110.34	6.7e+03	369	7082167	y939501.r1	Soares_fetal	gb_gss:AA553849	7.00	110.06	7.0e+03	382	AA553849	RPCT-23-177R20	TU R
gb_est2:BE8399465	7.00	110.34	6.7e+03	369	BE8399465	FC3-FN0143-020800-013		gb_gss:AA0298503	7.00	110.06	7.0e+03	382	AA0298503	HS_2244-A1-A03_MF C	
gb_est1:AI136740	7.00	110.31	6.7e+03	370	AI136740	rb23b01.y1	Zebrafish W	gb_est1:AA925889	7.00	110.03	7.0e+03	383	AA925889	265 P1RF62	Pinus ta
gb_est1:AJ346762	7.00	110.31	6.7e+03	370	AJ346762	SHR-000018-0-B16	HMS/S	gb_est1:AA925889	7.00	110.03	7.0e+03	383	AA925889	UI-R-A1-ec-h-08-0-U	
gb_est2:BU165460	7.00	110.31	6.7e+03	370	BU165460	BU165460	full length c	gb_est1:AV663660	7.00	110.03	7.0e+03	383	AV663660	AV663660	Bos taurus
gb_est2:BE577574	7.00	110.31	6.7e+03	370	BE577574	L48-2429M3	rice plant 1	gb_est1:AA362988	7.00	110.03	7.0e+03	383	AA362988	RCO-CR0300-11	1199-0
gb_est2:BE716508	7.00	110.31	6.7e+03	370	BE716508	CH1-HW0764-040700-299		gb_gss:AA0131186	7.00	110.03	7.0e+03	383	AA0131186	HS_3028-B1-E08_MF C	
gb_gss:AA048814	7.00	110.31	6.7e+03	370	AA048814	GSSBHU01270	Sheared gg	gb_est1:AI1642438	7.00	110.01	7.0e+03	384	AI1642438	vt04e09.x1	Bartstead
gb_est1:AA288043	7.00	110.29	6.8e+03	371	AA288043	va40G03.r1	Soares_mous	gb_est2:BI016321	7.00	110.01	7.0e+03	384	BI016321	PM3-ET0257-140301-0	
gb_est1:AV889043	7.00	110.29	6.8e+03	371	AV889043	AV889043	Nori Satoh un	gb_est1:AI1352289	7.00	109.99	7.0e+03	385	AI1352289	q12e01.x1	Soares_t
gb_est2:BG230233	7.00	110.29	6.8e+03	371	BG230233	uv04a07.x1	Soares_mous	gb_gss:AA0158271	7.00	109.99	7.0e+03	385	AA0158271	C00319-F	Lambda Zap
gb_est2:BI206493	7.00	110.29	6.8e+03	371	BI206493	EST524533	CRTOS Lycoper	gb_gss:AA0158366	7.00	109.99	7.0e+03	385	AA0158366	CITR-HSP-2312629	TF C
gb_gss:AA942822	7.00	110.29	6.8e+03	371	AA942822	2M020103P	Mouse 10Hd	gb_gss:AA851661	7.00	109.99	7.0e+03	385	AA851661	2M0154EB03	Mouse 10
gb_gss:AA0243362	7.00	110.29	6.8e+03	371	AA0243362	MS-2034-B2-D07	TY CIT	gb_est1:AV0002861	7.00	109.97	7.0e+03	386	AV0002861	AA002861	Mouse muscu
gb_est1:AA769597	7.00	110.27	6.8e+03	372	AA769597	mf7605a1.f1	Aspergillus	gb_est1:AA049378	7.00	109.97	7.0e+03	386	AA049378	EST903845	Mouse 7.5
gb_est1:AA488262	7.00	110.27	6.8e+03	372	AA488262	cx1-x10h07.r1	Soares_tote	gb_est2:BF8841192	7.00	109.97	7.0e+03	386	BF8841192	CM2-ET0193-191200-6	
gb_est2:BF882501	7.00	110.27	6.8e+03	372	BF882501	CM1-ET0197-051200-688		gb_gss:AA3645655	7.00	109.95	7.0e+03	387	AA3645655	CH230-195W2	TY CHOR
gb_est2:BEA06063	7.00	110.25	6.8e+03	373	BEA06063	WHE0408_f01	foliar whed	gb_gss:AA0840335	7.00	109.95	7.0e+03	387	AA0840335	F.rubripennis	GSS sequ
gb_est1:AA6299455	7.00	110.25	6.8e+03	373	AA6299455	h154f02.x1	Soares_NFL	gb_est1:AV0873779	7.00	109.95	7.0e+03	387	AV0873779	AV873779	NOIT Satoh
gb_est1:BB808553	7.00	110.25	6.8e+03	373	BB808553	BB808553	RIKEN full-16	gb_est2:BG318365	7.00	109.95	7.0e+03	387	BG318365	NXPEV_012_G06_F	NXPV
gb_est1:BB872699	7.00	110.25	6.8e+03	373	BB872699	BB872699	RIKEN full-16	gb_est2:BG318365	7.00	109.95	7.0e+03	387	BG318365	NXPEV_012_G06_F	NXPV
gb_est2:BI867210	7.00	110.25	6.8e+03	373	BI867210	f6c3h10.x1	Gong zebraf	gb_gss:AA0032707	7.00	109.95	7.0e+03	387	AA0032707	UI-M-GCOP-bny-A07_MF C	
gb_gss:AA0076403	7.00	110.25	6.8e+03	373	AA0076403	CITR-HSP-2359016	TY CIT	gb_gss:AA0631358	7.00	109.95	7.0e+03	387	AA0631358	RPCT-11-482B10	TY R
gb_gss:CNS00NLK	7.00	110.25	6.8e+03	373	AA082070	Arabidopsis thaliana g		gb_est1:AA9840335	7.00	109.95	7.0e+03	387	AA9840335	vc91d07.r1	Strateged
gb_est1:AM046273	7.00	110.23	6.8e+03	374	AM046273	UI-M-BH1-ala-f-05-0-U		gb_est1:AV004697	7.00	109.93	7.0e+03	388	AV004697	AV004697	Mus muscu
gb_est1:AM045721	7.00	110.23	6.8e+03	374	AM045721	15589 MARC lBOV Bos t		gb_est1:AM125651	7.00	109.93	7.0e+03	388	AM125651	UI-M-BH2-2-aq1-b-05	
gb_est1:AA4464637	7.00	110.23	6.8e+03	374	AA4464637	zx85b01.r1	Soares_ovar	gb_est1:AA512013	7.00	109.93	7.0e+03	388	AA512013	v142c08.r1	Strateged
gb_est1:BB809132	7.00	110.23	6.8e+03	374	BB809132	BB809132	RIKEN full-16	gb_est2:BG224239	7.00	109.93	7.0e+03	388	BG224239	1M003009	Bovine M1
gb_est2:RS91550	7.00	110.23	6.8e+03	374	RS91550	y910b07.s1	Soares_fet1	gb_gss:BB959600	7.00	109.93	7.0e+03	388	BB959600	BOHPISTF	BOH1 bras
gb_est2:3234455	7.00	110.23	6.8e+03	374	3234455	ATFS3279	Versallies-VR A	gb_est1:AA804010	7.00	109.91	7.0e+03	389	AA804010	GM12538	5Prtline GM D
gb_est2:BE455010	7.00	110.23	6.8e+03	374	BE455010	UI-M-B21-BKd-C-05-0-U		gb_est1:AA267435	7.00	109.91	7.0e+03	389	AA267435	m299B06.r1	Soares m
gb_gss:BBH342228	7.00	110.23	6.8e+03	374	BBH342228	CH230-42D17	TY CHOR1-2	gb_est1:AM814903	7.00	109.91	7.0e+03	389	AM814903	MRI-ST0206	-120400-0
gb_est1:AA975678	7.00	110.21	6.8e+03	375	AA975678	vv57d03.r1	Soares_thym	gb_est2:BI073010	7.00	109.91	7.0e+03	389	BI073010	C091P67U	Populus st
gb_est1:AA971464	7.00	110.21	6.8e+03	375	AA971464	op81d02.r1	Soares_NFL	gb_est2:BI073010	7.00	109.91	7.0e+03	389	BI073010	C091P67U	Populus st
gb_est2:BI976876	7.00	110.21	6.8e+03	375	BI976876	4866042	MARC 2BOV Bos t	gb_gss:BBH047513	7.00	109.91	7.0e+03	389	BBH047513	RPCT-24-346M17	TY R
gb_est2:BE994117	7.00	110.21	6.8e+03	375	BE994117	UI-M-GCOP-bdp-g-11-0-U		gb_est1:AA771940	7.00	109.89	7.0e+03	390	AA771940	a137c11.s1	Soares_P
gb_est1:AI047647	7.00	110.18	6.9e+03	376	AI047647	uh81a04.r1	Soares_mous	gb_est1:BB802699	7.00	109.89	7.0e+03	390	BB802699	BB802699	RIKEN full-1
gb_est2:BG431552	7.00	110.18	6.9e+03	376	BG431552	NFL110009D1F1076	Droeg	gb_est2:BE4279637	7.00	109.89	7.0e+03	390	BE4279637	MRO-GM0024-180800-0	
gb_est2:R42337	7.00	110.18	6.9e+03	376	R42337	yf91h02.s1	Soares_infant	gb_gss:AA0984000	7.00	109.89	7.0e+03	390	AA0984000	RPCT-23-266C23	TJB
gb_est2:728270	7.00	110.18	6.9e+03	376	728270	EST547747	Human Brain Hom	gb_gss:AA0984000	7.00	109.89	7.0e+03	390	AA0984000	RPCT-23-266C23	TJB
gb_est1:AA288555	7.00	110.16	6.9e+03	377	AA288555	m146903.r1	Life Tech m	gb_est1:AA475952	7.00	109.87	7.0e+03	391	AA475952	83213	MARC lBOV Bos
gb_est1:AAV547003	7.00	110.16	6.9e+03	377	AAV547003	AVS47003	Arabidopsis t	gb_est1:AA740811	7.00	109.87	7.0e+03	391	AA740811	uc71h01.x1	NCI_CGAP
gb_est1:BB803043	7.00	110.16	6.9e+03	377	BB803043	BB803043	RIKEN full-16	gb_est1:AA740811	7.00	109.87	7.0e+03	391	AA740811	uc71h01.x1	NCI_CGAP
gb_est2:BI9313124	7.00	110.16	6.9e+03	377	BI9313124	da187r02.y1	NICHD XGC	gb_gss:AA0945333	7.00	109.85	7.0e+03	392	AA0945333	ud41h12.r1	Soares_m
gb_est1:AI0194164	7.00	110.14	6.9e+03	378	AI0194164	ub21h03.r1	Soares_thym	gb_gss:AA0140756	7.00	109.85	7.0e+03	392	AA0140756	HS_3118-A1-A03_MF C	
gb_est1:AJ346806	7.00	110.14	6.9e+03	378	AJ346806	SHR-000018-0-D23	HMS/S	gb_est1:AI1664373	7.00	109.83	7.0e+03	393	AI1664373	ue61d04.r1	Soares_m
gb_est1:AA9492205	7.00	110.14	6.9e+03	378	AA9492205	UI-M-BH3-ctf-f-10-0-UI		gb_est1:AA4847109	7.00	109.83	7.0e+03	393	AA4847109	ab60904.r1	Strateged
gb_est2:BM281429	7.00	110.14	6.9e+03	378	BM281429	k133a01.y1	Ascaris suu	gb_gss:AA0124131	7.00	109.83	7.0e+03	393	AA0124131	HS_3112-A1-D09_MF C	
gb_gss:AA207914	7.00	110.14	6.9e+03	378	AA207914	RPCT-23-395D18	TU RPC1	gb_gss:AA0152055	7.00	109.83	7.0e+03	393	AA0152055	HS_3112-A1-B05_MF C	
gb_gss:BB291647	7.00	110.14	6.9e+03	379	BB291647	CH230-44B1	TU CHOR1-23	gb_est1:AA0843959	7.00	109.81	7.0e+03	394	AA0843959	zn07b11.r1	Strateged
gb_est1:AI1811243	7.00	110.12	6.9e+03	379	AI1811243	qv494a01.y1	Soares_mamm	gb_est1:AI1848325	7.00	109.81	7.0e+03	394	AI1848325	UI-M-AH1-ago-d-11-0	
gb_est1:AA8293317	7.00	110.12	6.9e+03	379	AA8293317	QV1-BM0001-301199-038		gb_est1:AA042521	7.00	109.81	7.0e+03	394	AA042521	AU042521	Mouse four
gb_gss:AA822348	7.00	110.12	6.9e+03	379	AA822348	qv418e2.x1	Ren Stubs</								

gb_est1.BB685206 + 7.00 109.78 7.2e+03 395 BB685206 BB685206 RIKEN full-16
 gb_est1.BE113612 + 7.00 109.78 7.2e+03 395 BE113612 UI-R-BJ1-awg-c-12-0-UI
 gb_est2.BE916806 - 7.00 109.78 7.2e+03 395 BE916806 IL3-up0115-141200-388-
 gb_est2.BG569258 - 7.00 109.78 7.2e+03 395 BG569254 TOSTT229805 v1 T9YEG1
 gb_est1.BB740658 + 7.00 109.76 7.2e+03 396 BB740658 BB740658 RIKEN full-16
 gb_gss:AA300716 + 7.00 109.74 7.3e+03 397 AA300716 HS-2202_AL-D02_MK CIT
 gb_gss:AA739945 + 7.00 109.72 7.3e+03 398 AA739945 PIUS taeda
 gb_est1.AA480457 - 7.00 109.72 7.3e+03 398 AA480457 ne10d04.s1 NCI CGAP EC
 gb_est1.AA504395 + 7.00 109.72 7.3e+03 398 AA504395 aa59d01.r1 NCI CGAP EC
 gb_gss:BB072891 + 7.00 109.72 7.3e+03 398 BB072891 RPCI-24-391Da.TJ RPCI-
 gb_est1:AV664980 + 7.00 109.70 7.3e+03 399 AV664980 AV664980 Bos taurus br
 gb_est1:AV732024 + 7.00 109.70 7.3e+03 399 AV732024 AV732024 HTF Homo sapi

seq_name: gb_est1:AU185039

seq_documentation_block:

LOCUS AU185039 143 bp mRNA linear EST 05-JUL-2001
 DEFINITION AU185039 Homo sapiens T-cell library (Sugita Y) Homo sapiens cDNA
 clone B00799-010, mRNA sequence.

ACCESSION AU185039
 VERSION AU185039.1 GI:14622952

KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 143)
 Sugita,Y., Oshida,T. and Oya,Y.
 TITLE Human cDNA sequencing
 JOURNAL Unpublished (2001)
 COMMENT Contact: Yuji Sugita
 Genex Research, Inc.
 907 Nogawa, Miyamae-ku, Kawasaki, Kanagawa 216-0001, Japan
 Tel.: 81-44-797-2281
 Fax: 81-44-797-2622
 Email: syuji@genex.co.jp, URL: http://www.genex.co.jp.

FEATURES
 source Location/Qualifiers
 1..143
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="B00799-010"
 /clone_lib="Homo sapiens T-cell library (Sugita Y)"
 /cell_type="T-cell"

BASE COUNT 48 a 31 c 27 g 37 t
 ORIGIN

alignment_scores:

Quality: 10.00 Length: 10
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AU185039 ..
 Align seg 1/1 to: AU185039 from: 1 to: 143

59 TTTATSerTherSerleuSerleuArgSer 68
 ||||||||||||||||||||||||||||
 18 TATGTTTCACCTCTTGTCTTCCTTAGGTCT 47

seq_name: gb_gss:AO528650

seq_documentation_block:

LOCUS AO528650 713 bp DNA linear GSS 18-MAY-1999
 DEFINITION RPCI-11-347C8.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-347C8,
 DNA sequence.

ACCESSION AO528650
 VERSION AO528650.1 GI:4840763

KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 713)
 Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and Venter

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 JOURNAL Unpublished (1997)
 COMMENT Other GSSs: RPCI-11-347C8-TV
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208

Email: hbe@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genet cs (info@resgen.com). BAC end search page:
 http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source Location/Qualifiers
 1..713
 /organism="Homo sapiens"
 /db_xref="GDB:7632919"
 /db_xref="taxon:9606"
 /clone="RPCI-11-347C8"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="vector: pBAC3.6; site.1: EcoRI; site.2: EcoRI;
 RPCI11 Human Male BAC library"

BASE COUNT 191 a 169 c 176 g 177 t
 ORIGIN

alignment_scores:

Quality: 10.00 Length: 10
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AO528650 ..
 Align seg 1/1 to: AO528650 from: 1 to: 713

62 ThrSerleuSerleuArgSerleuHisleu 71
 ||||||||||||||||||||||||||||
 497 ACTTCTTTGTCTTGAAGAGTGCACACCTC 526

seq_name: gb_est1:AV224462

seq_documentation_block:

LOCUS AV224462 231 bp mRNA linear EST 14-NOV-2001
 DEFINITION AV224462 RIKEN full-length enriched, 18 days pregnant, placenta and
 extra embryonic tissue Mus musculus cDNA clone 3830420N18 3', mRNA
 sequence.

ACCESSION AV224462
 VERSION AV224462.1 GI:6173840

KEYWORDS EST.
 SOURCE house mouse.

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 231)
 Komuro,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
 Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai
 C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,

Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigenome, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomioka, N., Tsunoda, Y., Watanabe, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Kono, H., et al. 1999)

Unpublished (1999)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel.: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.go.jp, genome.gsc.riken.go.jp, Saeki.N., Izawa.M., Watanabe.S., Muramatsu.M., Okazaki.Y. and Hayashizaki.Y.

Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)

Itch, M., Kitsuana, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details

TITLE
JOURNAL
COMMENT

```

      seq_name: gb_estl:AV292948
      seq_documentation_block:
        LOCUS       AV292948                     242 bp    mRNA    linear    EST 09-NOV-1996
        DEFINITION   AV292948 RIKEN full-length enriched, 6 days neonate head Mus
                    musculus cDNA clone 543043ID15 3', mRNA sequence.
        ACCESSION   AV292948
        VERSION     AV292948.1    GI:306979
        SOURCE      EST.
                   house mouse.
                   Mus musculus
        ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS          Kono,H., Alzawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
                  I (bases 1 to 242)
                  Fukuda,S., Fukuinishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Horii,F.,
                  Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Ka-
                  saki,Kawaji,T., Kikuchi,N., Kojoim,Y., Koya,S., Kusakabe,M.,
                  Matsuyama,T., Mikai,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
                  Owa,C., Ozawa,Y., Saito,H., Sano,M., Satou,K., Shibata,K., Shibeta-
                  su,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Suganari,Y., Suzuki,H.,
                  Suzuki,H., Takahashi,F., Tateno,M., Tomiyaga,N., Tsunoda,Y.,
                  Watabiki,A., Watanabe,S., Yamamuta,T., Yasunishi,A., Yokota,T.,
                  Yoshiki,A., Yoshino,M., Muramatsu,M., and Hayashizaki,Y.
TITLE            RIKEN Mouse ESTs (Kono,H., et al. 1999)
JOURNAL          Unpublished (1999)
COMMENT          Contact: yoshihide Hayashizaki

```

FEATURES
source

LOCUS	242 bp	mRNA	linear	EST 09-NOV-1999
DEFINITION	AV292948 RIKEN full-length enriched, 6 days neonate head Mus musculus cDNA clone 5430431D15 3', mRNA sequence.			
ACCESSION	AV292948			
VERSION	AV292948.1	GI:5306979		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 242)			
AUTHORS	Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kame,C., Kawai,T., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata,Y., Shigemoto,X., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tomiura,N., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamuta,T., Yasunishi,A., Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.			
TITLE	RIKEN Mouse ESTs (Konno,H., et al. 1999)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Yoshinide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Saitoh-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-resgsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Sasaki,N., Izawa,M., Watabiki,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y. Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998) Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Please visit our web site (http://genome.etc.riken.go.jp) for further details.			

seq_documentation_block: 242 bp mRNA linear EST 09-NOV-1999

LOCUS AV292948 RIKEN full-length enriched, 6 days neonate head Mus

DEFINITION AV292948 RIKEN full-length enriched, 6 days neonate head Mus

ACCESSION AV292948

VERSION AV292948

KEYWORDS AV292948.1 GI:6306979

ORGANISM EST.

SOURCE house mouse.

MUS musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 242)

AUTHORS Kono, H., Alzawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Isii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kasai, C., Kawaji, T., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Ode, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Suganari, Y., Suzuki, H., Suzuki, K., Takahashi, F., Tateo, M., Tomioka, N., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamuta, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Kono, H., et al. 1999)

COMMENT Unpublished (1999)

TITLE Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

CONTACT The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Saitama-shi, Tsukuba-shi, Ibaraki-shi, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.c.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Sasaki, N., Izawa, M., Watahiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)

Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Please visit our web site (<http://genome.etc.riken.go.jp>) for further details.

Location/Qualifiers

1..242

organism="Mus musculus"

strain="C57BL/6J"

db_xref="taxon:10090"

clone="5430431D15"

clone_lib="RIKEN full-length enriched, 6 days neonate head"

sex="mixed"

tissue_type="head"

dev_stage="6 days neonate"

lab_host="DH10B"

note="Site:1: Sali; Site:2: BamHI; CDNA library was

Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was


```

BBRC/SEERAD funded cereal IGF (Investigating Gene
Function) project."
BASE COUNT      100 a      170 c      96 g      77 t
ORIGIN
alignment_scores:
    Quality:      9.00      Length:      9
    Ratio:        1.000     Gaps:      0
    Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x BI776317      ..
Align seg 1/1 to: BI776317 from: 1 to: 443
      61 SerThrSerLeuArgSerAlaHisIleu 69
      |||||||
      89 TCCACCTCCCTTCCCTCCGCTCCGCGC 115
seq_name: gb_est1:AU168059
seq_documentation_block:
LOCUS      AU168059      727 bp      mRNA      linear      EST-29-JAN-2001
DEFINITION      AU168059 Ol-br-ad cDNA Oryzias latipes cDNA clone br1636, mRNA
sequence.
ACCESSION      AU168059
VERSION        AU168059.1 GI:12590128
KEYWORDS       EST.
SOURCE         Japanese medaka.
ORGANISM       Oryzias latipes.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorphae; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 727)
Mita, K., Ishikawa, Y., and Yamuchi, M.
Establishment of cDNA database of medaka, Oryzias latipes
Unpublished (2001)
Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmita@nirs.go.jp
method: uni-directional sequence direction: sequenced from T3 primer
(5' -> 3').
FEATURES
    source
        location/Qualifiers
            1..727
                /organism="Oryzias latipes"
                /strain="HNI"
                /db_xref="taxon:8090"
                /clone="br1636"
                /clone_1bp="Ol-br-ad cDNA"
                /sex="female/male mixed"
                /tissue_type="brain"
                /dev_stage="adult"
BASE COUNT      165 a      211 c      217 g      132 t      2 others
ORIGIN
alignment_scores:
    Quality:      9.00      Length:      9
    Ratio:        1.000     Gaps:      0
    Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AU168059/rev      ..
Align seg 1/1 to reverse of: AU168059 from: 1 to: 727
      63 SerLeuSerLeuArgSerAlaHisIleu 71
      |||||||
      504 AGTCGTGATTCAGGTGACGCTCATCTC 478

```


seq_name: gb_gss:A0274510
seq_documentation_block:
LOCUS A0274510 768 bp DNA linear GSS 03-NOV-1998
DEFINITION mgxb0022a11r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgxb0022a11r, DNA sequence.
ACCESSION A0274510
VERSION A0274510
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
1 (bases 1 to 768)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasnowski, M., Wang, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
TITLE Unpublished (1998)
JOURNAL Contact: Dean RA
COMMENT Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 472.
Location/Qualifiers
1..768
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0022a11r"
/clone_1lb="CUGI Rice Blast BAC Library"
/lssue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice Blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."

BASE COUNT 149 a 210 c 241 g 167 t 1 others
ORIGIN

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
us-09-528-682-1 x A0274510 ..
Align seg 1/1 to: A0274510 from: 1 to: 768

seq_name: gb_gss:BH111124
seq_documentation_block:
LOCUS BH111124 104 bp DNA linear GSS 19-JUL-2001
DEFINITION RPCI-24-367M2.TJ RPCI-24 Mus musculus genomic clone RPCI-24-367M2,

ACCESSION BH111124
VERSION BH111124.1 GI:14945954
KEYWORDS GSS.
SOURCE Mus mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 104)
AUTHORS Zhao, S., Nierman, W., Malek, J., Shatman, S., Akhurst, B., Levins, M.,
Tsegaye, G., Geer, K., Kroll, M., Shwartsbeyn, A., Gebregeorgis, E.,
Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSSs: RPCI-24-367M2.TJ
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (<http://www.choi.tigr.org/bacpac/orderingframe.html>). BAC end
page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 367 Row: M Column: 2
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..104
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-367M2"
/clone_1lb="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTRABAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTRABAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT 17 a 24 c 35 g 28 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
us-09-528-682-1 x BH111124 ..
Align seg 1/1 to: BH111124 from: 1 to: 104

seq_name: gb_est2:BI778862
seq_documentation_block:
LOCUS BI778862 137 bp mRNA linear EST 26-SEP-2001
DEFINITION EBR001_S0001_M17_R IGF Barley EBR001 library Hordeum vulgare cDNA
clone EBR001_S0001_M17 5', mRNA sequence.
ACCESSION BI778862
VERSION BI778862.1 GI:15781754
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Bakaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae ; Trilicaceae; Hordeum.
1 (bases 1 to 137)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L., Ramsey, L., Machray, G., Marshall, D.F.M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R
Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.

FEATURES
Source
Location/Qualifiers

1..137
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBR001.S0001.M17"
/clone_lib="IGF Barley EBR001 library"
/tissue_type="Root, unstressed"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site.1: Sal I; Site.2: Not I; Non-normalised library, directionally cloned into pSPORT1. Derived from roots of 3 week old hydroponically grown unstressed barley plants. Developed as part of the Barley Transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating Gene Function) project."

BASE COUNT
ORIGIN

46 a 35 c 38 g 18 t

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x B1778862/rev ..

Align seg 1/1 to reverse of: B1778862 from: 1 to: 137

61 SerThrSerLeuSerLeuArgSer 68
|||||
99 AGCACCCTCTCTCTCTCAGATCT 76

seq_name: gb_est1:BB550100

seq_documentation_block:
LOCUS BB550100 141 bp mRNA linear EST 01-AUG-2000
DEFINITION BB550100 RIKEN full-length enriched, 2 days pregnant adult female
OVIDUCT Mus musculus cDNA clone E230020G08 3', mRNA sequence.
VERSION BB550100
KEYWORDS BB550100.1 GI:9636466
EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 141)
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hata, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Itawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Maki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya

TITLE
JOURNAL
COMMENT

Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermotabilization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitsuai, T., Akiyama, Y., Shibata, K., Itawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
Source
Location/Qualifiers

1..141
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E230020G08"
/clone_lib="RIKEN full-length enriched, 2 days pregnant adult female oviduct"
/sex="Female"
/tissue_type="oviduct"
/dev_stage="2 days pregnant adult"
/lab_host="DH10B"
/note="Site.1: SalI; Site.2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCCGACCTGAGTTTCTTTTCTTTT 3'], cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT
ORIGIN

25 a 74 c 10 g 32 t

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BB550100 ..

Align seg 1/1 to: BB550100 from: 1 to: 141

64 LeuSerLeuArgSerAlaHisLeu 71
|||||
4 CTCCTACTCCGTTCCGCCACCTC 27


```

seq_name: gb_est2:BF292752

seq_documentation_block:
LOCUS       BF292752             159 bp    mRNA             linear   EST 17-NOV-2000
DEFINITION  WHE2202.D09_G182S Aegilops speltoides anther cDNA library Aegilops
ACCESSION   BF292752
VERSION     BF292752.1
KEYWORDS    GI:11223816
SOURCE      Aegilops speltoides.
ORGANISM    Aegilops speltoides.
REFERENCE   1 (bases 1 to 159)
AUTHORS     Akhunov, E., Anderson, O.D., Chao, S., Chin, A., Choi, D.W., Close, T.J.,
            Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Kianian, P., Lazo, G.R.,
            Miller, R., Otto, C., Rausch, C.J., Seaton, C.L., Simons, K., Tong, J.C.
            and Zhang, D.
TITLE       The structure and function of the expressed portion of the wheat
COMMENT     genomes - Anther cDNA library from Aegilops speltoides
JOURNAL     Unpublished (2000)
CONTACT     Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105395818
            Email: cander@nps.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: StrataGene SK primer.
            Location/Qualifiers
                1..159
                /organism="Aegilops speltoides"
                /cultivar="F2 from 2-12-4-8-1-1-(1) x P136909-12-811-(1)"
                /db_xref="taxon:4573"
                /clone_1lb="WHE2202.D09_G18"
                /clone_1lb="Aegilops speltoides anther cDNA library"
                /tissue_type="Anther"
                /dev_stage="Premelotic anthers"
                /lab_host="E. coli SOLR"
                /note="Vector: Lambda Uni-ZAP XR, excised phagemid:
                Site_1: EcoRI, Site_2: XhoI; Plants were grown in a growth
                chamber at the University of California, Davis (Akhunov).
                Premelotic anthers were harvested, total RNA and poly(A)
                RNA were prepared, from each tissue and then pooled, a
                cDNA library was made, and the cDNA clones were in vivo
                excised to give phagescript phagemids in the Tj Close lab
                (Akhunov, Chin, Choi, Close, Fenton, Kianian, Otto, Simons
                , Zhang) at the University of California, Riverside.
                Plasmid DNA preparations and DNA sequencing were performed
                in the OD Anderson lab (all other authors)."
```

```

BASE COUNT      34 a      29 c      45 g      51 t
ORIGIN
```

```

alignment_scores:
    Quality:      8.00      Length:      8
    Ratio:        1.000     Gaps:        0
Percent Similarity: 100.000    Percent Identity: 100.000
```

```
alignment_block:
```

```
US-09-528-682-1 x BF292752 ..
```

```
Align seg 1/1 to: BF292752 from: 1 to: 159
```

```

61 SerThrSerLeuSerLeuArgSer 68
|||||
34 TCGACAAAGCTTCTCTCGCAAGT 57
seq_name: gb_est2:BE517249
```

```

seq_documentation_block:
LOCUS       BE517249             163 bp    mRNA             linear   EST 08-AUG-2000
DEFINITION  WHE0616.G08_N162A wheat ABA-treated embryo cDNA library Triticum
ACCESSION   BE517249
VERSION     BE517249.1
KEYWORDS    GI:9741279
SOURCE      bread wheat.
ORGANISM    Triticum aestivum
REFERENCE   1 (bases 1 to 163)
AUTHORS     Anderson, O.D., Chao, S., Han, P.S., Hsia, C.C., Johnson, R.R., Kang, Y.,
            Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L., Tong, J.C., Verhey
            , S.D. and Walker-Simmons, M.K.
TITLE       The structure and function of the expressed portion of the wheat
COMMENT     genomes - ABA-treated embryo library
JOURNAL     Unpublished (2000)
CONTACT     Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: cander@nps.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Clontech Matchmaker 3' AD primer.
            Location/Qualifiers
                1..163
                /organism="Triticum aestivum"
                /cultivar="Brevor (soft, white, winter, common wheat)"
                /db_xref="taxon:4565"
                /clone_1lb="WHE0616.G08_N16"
                /clone_1lb="Wheat ABA-treated embryo cDNA library"
                /tissue_type="Seed dormant"
                /dev_stage="mature dormant seeds"
                /lab_host="E. coli DH12S"
                /note="Vector: pGAD10; Site_1: EcoRI; Site_2: XhoI;
                Embryos were cut from mature, dormant seeds and imbedded in
                25 microm ABA (abscisic acid) in 5 mm MMS buffer, pH 5.7,
                for 12 hr at 22 C. The tissue, total RNA, and poly(A)
                RNA were prepared by Steven Verhey in M.K.
                Walker-Simmons' lab (USDA-ARS, Washington State Univ.,
                Pullman, Washington 99164-6420). A cDNA library was made
                by Clontech using a combination of random and oligo dT
                primers. Library was plated and archived by Russell
                Johnson (Colby College, ME/Walker-Simmons' lab). Plasmid
                DNA preparations and DNA sequencing were performed in the
                OD Anderson lab (all other authors)."
```

```

BASE COUNT      39 a      36 c      54 g      34 t
ORIGIN
```

```

alignment_scores:
    Quality:      8.00      Length:      8
    Ratio:        1.000     Gaps:        0
Percent Similarity: 100.000    Percent Identity: 100.000
```

```
alignment_block:
```

```
US-09-528-682-1 x BE517249 ..
```

```
Align seg 1/1 to: BE517249 from: 1 to: 163
```

```

61 SerThrSerLeuSerLeuArgSer 68
|||||
93 TCGACAAAGCTTCTCTCGCAAGT 116
seq_name: gb_est1:AA637879
seq_documentation_block:
```


LOCUS AA637879 169 bp mRNA linear EST 22-OCT-1997
 DEFINITION v129h03.r1 Barstead mouse myotubes MRLB5 mus musculus cDNA clone
 IMAGE:1122101 5' similar to TR:G1020151 G1020151 CLP36. ;, mRNA
 sequence.
 ACCESSION AA637879.1 GI:2561467
 VERSION AA637879.1
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 169)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Gettel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:611437
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28m13 rev2 EF from Amersham
 High quality sequence stop: 1.
 FEATURES
 source
 1..169
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C3H"
 /db_xref="taxon:10090"
 /clone="IMAGE:1122101"
 /clone_lib="Barstead mouse myotubes MRLB5"
 /cell_line="C2C12"
 /lab_host="DH10B"
 /note="Vector: p7T73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTTCGATCGATGAGTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 [AATTCGATCCTTG], digested with Not I and cloned into the
 Not I and Eco RI sites of the modified p7T73 vector.
 Library constructed by Bob Barstead. The C2C12 cell line
 (available from ATCC, catalog # CRL-1772) differentiates
 rapidly, forming contractile myotubes and producing
 characteristic muscle proteins."
 BASE COUNT 30 a 49 c 45 g 45 t
 ORIGIN
 alignment_scores: Length: 8
 Quality: 1.000 Gaps: 0
 Ratio: 1.000 Percent Identity: 100.000
 Percent Similarity: 100.000
 alignment_block:
 US-09-528-682-1 x AA637879 ..
 Align seg 1/1 to: AA637879 from: 1 to: 169
 62 ThrSerLeuSerLeuArgSerAla 69
 ||||||||||||||||||||
 51 ACTTCACCTCAGCGCTCGGCTCAGCA 74
 seq_name: gb_est2:BM373106

seq_documentation_block:
 LOCUS BM373106 232 bp mRNA linear EST 10-JAN-2002
 DEFINITION EBMa04_S0003_C02_R IGF Barley EBMa04 library Hordeum vulgare cDNA
 clone EBMa04_S0003_C02 5', mRNA sequence.
 ACCESSION BM373106
 VERSION BM373106.1
 KEYWORDS EST.
 SOURCE barley.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 232)
 AUTHORS Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardie, L.,
 Ramsay, L., Machray, G., Marshall, D.F.M. and Maugh, R.
 TITLE Development of Barley Transcriptome Resources
 JOURNAL Unpublished (2001)
 COMMENT Contact: Maugh R
 Unit of Genomics
 Scottish Crop Research Institute
 Invergowrie, Dundee, DD2 5DA, Scotland, UK
 Tel: 00 44 1382 562731
 Fax: 00 44 1382 562426
 Email: rwaugh@scri.sari.ac.uk
 All sequence has a Phred quality score of 20 or over
 Seq primer: M13 reverse.
 FEATURES
 source
 1..232
 Location/Qualifiers
 /organism="Hordeum vulgare"
 /cultivar="Optic"
 /db_xref="taxon:4513"
 /clone="EBMa04_S0003_C02"
 /clone_lib="IGF Barley EBMa04 library"
 /tissue_type="Maternal tissue"
 /dev_stage="10 days post anthesis"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
 Non-normalised library, directionally cloned into pSPORT1.
 Derived from maternal tissue dissected from developing
 grains (10 days post anthesis) in glasshouse grown barley
 plants. Developed as part of the barley transcriptome
 resources of BBSRC/NERAC funded cereal IGF (Investigating
 Gene Function) project."
 BASE COUNT 32 a 82 c 77 g 41 t
 ORIGIN
 alignment_scores: Length: 8
 Quality: 1.000 Gaps: 0
 Ratio: 1.000 Percent Identity: 100.000
 Percent Similarity: 100.000
 alignment_block:
 US-09-528-682-1 x BM373106/rev ..
 Align seg 1/1 to reverse of: BM373106 from: 1 to: 232
 61 SerThrSerLeuSerLeuArgSer 68
 ||||||||||||||||||||
 45 AGCACCTCTCTCTCTCAGATCT 22
 seq_name: gb_est1:AW202162
 seq_documentation_block:
 LOCUS AW202162 254 bp mRNA linear EST 02-DEC-2001
 DEFINITION s112f04.y1 Gm-c1027 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-c1027-2072 5' similar to TR:Q92V18 Q92V18 T6A23.14 PROTEIN. ;,
 mRNA sequence.
 ACCESSION AW202162
 VERSION AW202162.1
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 254)
Shoemaker, R., Kelm, F., Vodka, L., Ertelding, J., Coryell, V., Khanna

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 254)	Shoemaker, R., Kaim, P., Vodka, L., Eppelding, J., Coryell, V., Khana, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepte, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.	Public soybean EST project	Unpublished (1999)	Contact: Shoemaker R/Public Soybean EST project

This clone is available through: Reegen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cu@reegen.com
Insert Length: 492 Std Error: 0.00.

FEATURES

Source

1. .254
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1027-2072"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/note="Vector: pBluescript II SK(-), Site.1: EcoRI, Site.2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First-stranded synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer [GAGAGGAGAGAGAGAGAGAACTACTCTCGAG(7)18] to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in
with cloned Pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The XhoI sites within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500 bp cutoff,
using GIBCOBRL Life Technologies' cDNA Size Fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript II(-) XR predigested vector
(pBluescript II SK(-) that has been digested with EcoRI
and XhoI, and phosphorylated by Stratagene). 97% of the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=30). This
library was constructed by Dr. Paul Kelm and Dr. Virginia
Corryell."

BASE COUNT	60 a	58 c	70 g	66 t
ORIGIN				

ORIGIN

```

alignment_scores:
  Quality: 8.00      Length: 8
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment_block:
US-09-528-682-1 x AW202162

Align seg 1/1 to: AW202162 from: 1 to: 254

159 GluaspGlyTyrArgLeuAlaGly 166
 |||||
 169 GAGGATGGTTATCGCCCTCGCAGGT 192

seq_name: gb_est1:AA998803

```
seq_documentation_block:
```

LOCUS	AA9596803	258 bp	linear	EST 04-JUL-1999
DEFINITION	UI-R-CO-lm-h-12-0-UI s1	UI-R-CO	Rattus norvegicus	cdna clone
UI-R-CO-lm-h-12-0-UI 3'				mRNA sequence.

VERSION	AA998803.1	GI:4292381
KEYWORDS	EST.	
SOURCE	Norway rat.	
ORGANISM	Rattus norvegicus	

Eukaryota; Metazoa; Choriata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE	1 (bases 1 to 258)
AUTHORS	Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
COMMENT	On Jun 5, 1998 this sequence version replaced gi:3189454.

Email: msosares@blue.weeg.uiowa.edu
The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized adult 8-day-Embryo library. cDNA Library Preparation: M. Fatima Bonaldo, Ph.D. Clone distribution: clones will be available through Research Genetics This clone is also available through the I.M.A.G.E. Consortium at LNL (info@image.llnl.gov). IMAGE ID=1782402

FEATURES

Source

Location/Qualifiers

1. 258

organism="Rattus norvegicus"
strain="Sprague-Dawley"
db_xref="taxon:10116"
clone="UI-R-C0-1m-h-12-0-UI"
clone_1fb="UI-R-C0"
dev_stage="adult"
lab_host="DH10B (Life Technologies)"
note="Vector: pT730-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI. The UI-R-C0
library is a subtracted library derived from the UI-R-A1
and UI-R-E1 libraries. The UI-R-A1 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, and muscle. The UI-R-E1
library consisted of a mixture of individually tagged
normalized libraries constructed from 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which
allows identification of the library of origin of a clone
within the mixture. The subtracted library (UI-R-C0) was
constructed as follows: PCR amplified cDNA inserts from a
pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had
been derived was used as a driver in a hybridization with
the pooled UI-R-A1 and UI-R-E1 library in the form of
single-stranded circles. The remaining single-stranded

circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-CO library. This procedure has been previously described (Bonald, Lennon and Soares, Genome Research 6: 791-806, 1996).

BASE COUNT 66 a 50 c 66 g 76 t
ORIGIN

alignment_scores:
quality: 8.00 length: 8
ratio: 1.000 gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AA998803/rev ..

Align seg 1/1 to reverse of: AA998803 from: 1 to: 258

60 ValSerThrSerLeuSerLeuArg 67
|||||
152 GTATCTACTCTCTCTCTTAAGG 129

seq_name: gb_gss:A2788524

seq_documentation_block:

LOCUS A2788524 263 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0035K08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0035K08 R, DNA sequence.

ACCESSION A2788524
VERSION A2788524.1 GI:12928413

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 263)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

TITLE Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

JOURNAL
COMMENT

Insert Length: 10000 Std Error: 0.00
Plate: 0035 row: K column: 08
Seq primer: CACACAGCAACACGATGACC
Class: plasmid ends
High quality sequence stop: 263.
Location/Qualifiers

FEATURES
source

1.263
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0035K08"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.."

BASE COUNT 47 a 81 c 78 g 57 t
ORIGIN

alignment_scores:
quality: 8.00 length: 8
ratio: 1.000 gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A2788524 ..

Align seg 1/1 to: A2788524 from: 1 to: 263

69 AAlaHSteuAlaGlyGlnSerIle 76
|||||
161 GCTCATCTGCGGGAGACGACATT 184

seq_name: gb_gss:AQ911978

seq_documentation_block:

LOCUS AQ911978 269 bp DNA linear GSS 25-MAY-2001
DEFINITION LMAJFV1_In09B09.y1 Leishmania major FV1 random genomic library
Leishmania major genomic clone LMAJFV1_In09B09 5', DNA sequence.

ACCESSION AQ911978
VERSION AQ911978.1 GI:6508494

KEYWORDS GSS.

SOURCE Leishmania major.

ORGANISM Leishmania major

Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania.

REFERENCE 1 (bases 1 to 269)
Akopyants, N.S., Clifton, S.W., Martin, J., Pape, D., Wylie, T., Li, L.,
Kissinger, J., Roos, D.S., Marra, M., Hillier, L., Chinwalla, A.,
Bishtain, A., Schmitt, A., Person, B., Theising, B., Ritter, E., Ronko,
I., Bennett, J., Cole, R., Underwood, R., Cardenas, M., Gibbons, M.,
Harvey, N., McCann, R., Tsagaris, R., Williams, T., Jackson, Y.,
Bowers, Y., Swaller, T., Waterston, R., Wilson, R. and Beverley, S.M.

TITLE

JOURNAL
MEDLINE
COMMENT
Other GSSs: In09B09.x1
Contact: Akopyants, NS / Beverley, SM
Washu Leishmania Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

Library construction: Natalia S. Akopyants, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
If using this information please cite:
N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major
Friedlin strain Y1 genome by shotgun sequencing' and the Washington
University Genome Sequencing Center for information on obtaining
clone material please contact: Natalia S. Akopyants Ph.D.
(natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.

(beverley@borcim.wustl.edu)
 Seq primer: -40RP from Gibco
 Class: Shotgun
 High quality sequence stop: 248.
 Location/Qualifiers
 1..269

BASE COUNT 52 a 80 c 59 g 78 t
 ORIGIN
 alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AG911978 ..

Align seg 1/1 to: AG911978 from: 1 to: 269

61 SerThSerLeuSerLeuArgSer 68
 |||
 78 TCCACATCTCTTCATTCGCTTCC 101

seq_name: gb_est2:BF908771

seq_documentation_block:

LOCUS BF908771 275 bp mRNA linear EST 18-JAN-2001
 DEFINITION OVI-UT00096-121000-418-b06 UT0096 Homo sapiens CDNA, mRNA sequence.
 ACCESSION BF908771
 VERSION BF908771.1 GI:12300229
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 275)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 'M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-3704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl-QV1st2-QV1-UT0096-
 121000-418-b06&cl=2000-10-12&cl=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 234.

FEATURES
 source
 Location/Qualifiers
 1..275

BASE COUNT 75 a 47 c 44 g 109 t
 ORIGIN
 alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x BF908771/rev ..

Align seg 1/1 to reverse of: BF908771 from: 1 to: 275

63 SerLeuSerLeuArgSerAlaHis 70
 |||
 41 TCATTATCTTACGAGTGCACAC 18

seq_name: gb_est2:W90748

seq_documentation_block:

LOCUS W90748 309 bp mRNA linear EST 07-MAY-1997
 DEFINITION zh79f01.s1 Soares_fetal_liver_spleen_infls.S1 Homo sapiens CDNA
 clone IMAGE:418297 3', mRNA sequence.
 ACCESSION W90748
 VERSION W90748.1 GI:1406714
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 309)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
 Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins
 'M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore
 'B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
 Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevaekis, E.,
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 97044478

CONTACT: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu

This clone is available royalty-free through LIND; contact the
 IMAGE Consortium (info@image.jnl.gov) for further information.
 Insert length: 947 Std Error: 0.00
 Seq primer: mob.REGA+ET
 High quality sequence stop: 283.

FEATURES
 source

Location/Qualifiers
 1..309
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="418297"
 /clone_image="418297"
 /clone_image="418297"
 /sex="male"


```

/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/Note="Organ: Liver and Spleen; Vector: pT77rd (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5',
AAGTCGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldi."
BASE COUNT      100 a      61 c      53 g      94 t      1 others
ORIGIN

alignment_scores:
    Quality:      8.00      Length:      8
    Ratio:        1.000      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x W90748/rev ..

Align seg 1/1 to reverse of: W90748 from: 1 to: 309

220 ArgGlnIlePheSerAspTyrGln 227
|||||
69 AGACAAATCTTCTGATTCATCA 46

seq_name: gb_est2:BI050522

seq_documentation_block:
LOCUS      BI050522      312 bp      mRNA      linear      EST 15-JUN-2001
DEFINITION PM2-GN0015-060101-008-f04 GN0015 Homo sapiens CDNA, mRNA sequence.
ACCESSION  BI050522
VERSION     BI050522.1 GI:14458052
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 312)
            Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
            Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
            , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.G.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM2&t2=PM2-GN0015-
060101-008-f04&t3=2001-01-06&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 18
Location/Qualifiers
1..312
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source

```

```

/clone_id="GN0015"
/dev_stage="Adult"
/Note="Organ: Placenta normal; Vector: puc18; Site 1: Sma I
; Site 2: Sma I; A mini-library was made by cloning
products derived from ONESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT      58 a      98 c      87 g      69 t
ORIGIN

alignment_scores:
    Quality:      8.00      Length:      8
    Ratio:        1.000      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BI050522/rev ..

Align seg 1/1 to reverse of: BI050522 from: 1 to: 312

111 GlnGluValSerAlaLeuGlyGly 118
|||||
176 CAGAGAGTCACAGACTGGGCGC 153

seq_name: gb_est2:BI781014

seq_documentation_block:
LOCUS      BI781014      312 bp      mRNA      linear      EST 26-SEP-2001
DEFINITION Ebma03_S0001.C14.R IGF Barley Ebma03 library Hordeum vulgare cDNA
clone Ebma03_S0001.C14.5', mRNA sequence.
ACCESSION  BI781014
VERSION     BI781014.1 GI:15783906
KEYWORDS   EST.
SOURCE     barley.
ORGANISM   Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
            ; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 312)
            Hedley, P., Liu, H., Caldwell, D., McCallum, N., Kudrie, S., Cardie, L.,
            Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
            Development of Barley Transcriptome Resources
            Unpublished (2001)
            Contact: Waugh R
            Unit of Genomics
            Scottish Crop Research Institute
            Invergowrie, Dundee, DD2 5DA, Scotland, UK
            Tel: 00 44 1382 562731
            Fax: 00 44 1382 562426
            Email: rwaugh@scri.sari.ac.uk
            All sequence has a Phred quality score of 20 or over
            Seq primer: M13 reverse.
            Location/Qualifiers
            1..312
            /organism="Hordeum vulgare"
            /cultivar="Optic"
            /db_xref="taxon:4513"
            /clone="Ebma03_S0001.C14"
            /clone_id="IGF Barley Ebma03 library"
            /tissue_type="Maternal tissue"
            /dev_stage="8 days post anthesis"
            /lab_host="DH10B"
            /Note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
            Non-normalised library, directionally cloned into pSPORT1.
            Derived from maternal tissue dissected from developing
            grains (8 days post anthesis) in glasshouse grown barley
            plants. Developed as part of the barley transcriptome
            resources of BBSRC/SEERAD funded cereal IGF (Investigating
            Gene Function) project."
            Gene function
            83 a      54 c      88 g      87 t
BASE COUNT

```


ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x B1781014 ..

Align seg 1/1 to: B1781014 from: 1 to: 312

61 SerTherSerLeuSerLeuArgSer 68
|||||
77 TCGACAGCTTGTCTCTCGCAGAGT 100

seq_name: gb_est2:B178101253

seq_documentation_block:

LOCUS B178101253 320 bp mRNA linear EST 21-NOV-2001
DEFINITION EBP101_SQ003_F08.R IGF Barley EBP101 library Hordeum vulgare cDNA
clone EBP101_SQ003_F08 5', mRNA sequence.

ACCESSION B178101253

VERSION B178101253.1 GI:17032321

KEYWORDS EST.

SOURCE barley.

ORGANISM Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 320)

REFERENCE Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,

Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.

Development of Barley Transcriptome Resources

Unpublished (2001)

JOURNAL

COMMENT

Contact: Waugh R

Unit of Genomics

Scottish Crop Research Institute

Invergowrie, Dundee, DD2 5DA, Scotland, UK

Tel: 00 44 1382 562731

Fax: 00 44 1382 562426

Email: rwaugh@scri.sari.ac.uk

All sequence has a Phred quality score of 20 or over

Seq primer: M13 reverse

FEATURES

Location/Qualifiers

1..320

/organism="Hordeum vulgare"

/cultivar="Optic"

/db_xref="taxon:4513"

/clone="EBP101_SQ003_F08"

/clone_1lb="IGF Barley EBP101 library"

/tissue_type="Pistils"

/dev_stage="24 hours post anthesis"

/lab_host="DH10B"

/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;

Non-normalised library, directionally cloned into pSPORT1.

Derived from pistils dissected from developing grains (24

hours post anthesis) in glasshouse grown barley plants.

Developed as part of the barley transcriptome resources of

BBSRC/SEERAD funded cereal IGF (Investigating Gene

Function) project."

BASE COUNT

ORIGIN

68 a 57 c 98 g 97 t

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x B178101253 ..

Align seg 1/1 to: B178101253 from: 1 to: 320

61 SerTherSerLeuSerLeuArgSer 68

|||||
84 TCGACAGCTTGTCTCTCGCAGAGT 107

seq_name: gb_est2:BF778513

seq_documentation_block:

LOCUS BF778513 323 bp mRNA linear EST 12-JAN-2001
DEFINITION NXSI_085_F03_F NXSI (NsI Xylem Side wood Inclined) Pinus taeda cDNA
clone NXSI_085_F03 5', mRNA sequence.

ACCESSION BF778513

VERSION BF778513.1 GI:12126413

KEYWORDS EST.

SOURCE 10b101ly pine.

ORGANISM Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 323)

REFERENCE Sederoff, R.

Molecular Basis of Wood Formation in the Pine Megagenome

Unpublished (2000)

JOURNAL

COMMENT

Contact: Johnson, Archur

North Carolina State University

Tel: 919 515 7800

Fax: 919 515 7801

Email: ajohnson@unity.ncsu.edu

Seq primer: 73.

FEATURES

Location/Qualifiers

1..323

/organism="Pinus taeda"

/strain="Coastal plain loblolly pine from North Carolina"

/db_xref="taxon:3352"

/clone="NXSI_085_F03"

/clone_1lb="NXSI (NsI Xylem Side wood Inclined)"

/tissue_type="Xylem"

/cell_type="Side"

/dev_stage="Juvenile"

/lab_host="XL1-Blue"

/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI

The library is from early (spring) wood, taken from

three six-year old trees (three different genotypes), in

the juvenile phase. These trees were induced to form side

wood by bending to a 45 degree angle and tying them to the

ground. Differentiating xylem was harvested from the sides

of the inclined stems, and a mixture of all three

genotypes was used for the library. oligo-dT primed cDNA

was directionally cloned into the EcoRI-XhoI Bluescript SK

vector arms. NOTE: The sequences contain a 'cDNA adapter'

between the EcoRI site and the start of the EST. The

adapter sequence is 'ATATCGGACGAG'."

BASE COUNT

ORIGIN

59 a 84 c 82 g 89 t 9 others

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x BF778513 ..

Align seg 1/1 to: BF778513 from: 1 to: 323

14 AspGluIleLysArgSerGly 21

|||||

216 GATGAGATCAAGCGGAGCGG 239

seq_name: gb_est2:BM368647


```

seq_documentation_block:
LOCUS      BM368647          333 bp      mRNA      linear      EST 10-JAN-2002
DEFINITION EBem08_SQ004_D10_R IGF Barley EBem08 library Hordeum vulgare cDNA
ACCESSION  BM368647
VERSION     BM368647.1  GI:18112037
KEYWORDS   EST.
SOURCE     Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae
            ORIGIN
            1..333
            /organism="Hordeum vulgare"
            /cultivar="Optic"
            /db_xref="taxon:4513"
            /clone="EBem08_SQ004_D10"
            /clone_lib="IGF Barley EBem08 library"
            /tissue_type="Embryo"
            /dev_stage="40 days post anthesis"
            /lab_host="DH10B"
            /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
            Non-normalised library, directionally cloned into pSPORT1.
            Derived from embryos dissected from developing grains (40
            days post anthesis) in glasshouse grown barley plants.
            Developed as part of the barley transcriptome resources of
            BBSRC/SEERD funded cereal IGF (Investigating Gene
            Function) project."
            FUNCTION project.
BASE COUNT      80 a      58 c      99 g      96 t
ORIGIN
alignment_scores:
      Quality:      8.00      Length:      8
      Ratio:      1.000      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x BM368647 ..
Align seg 1/1 to: BM368647 from: 1 to: 333
61 SerThrsSerLeuSerLeuArgSer 68
|||||
109 TCACACAAAGCTTCTCTCGCAAGT 132
seq_name: gb_gss:AZ094117
seq_documentation_block:
LOCUS      AZ094117          346 bp      DNA      linear      GSS 08-MAY-2000
DEFINITION RPCI-23-448B16.T1 RPCI-23 Mus musculus genomic clone RPCI-23-448B16
ACCESSION  AZ094117
VERSION     AZ094117.1  GI:7736160
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 346)
            Zhao,S., Niemman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinet
            B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
            Mouse BAC End Sequences from Library RPCI-23
            Unpublished (1999)
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 448 row: B column: 16
            Seq primer: SP6
            Class: BAC ends.
FEATURES
SOURCE
1..346
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-448B16"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and, partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBAC3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      94 a      70 c      97 g      85 t
ORIGIN
alignment_scores:
      Quality:      8.00      Length:      8
      Ratio:      1.000      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x AZ094117/rev ..
Align seg 1/1 to reverse of: AZ094117 from: 1 to: 346
61 SerThrsSerLeuSerLeuArgSer 68
|||||
104 AGCACCCTCTCTAGCGCTCC 81
seq_name: gb_gss:AQ123154
seq_documentation_block:
LOCUS      AQ123154          355 bp      DNA      linear      GSS 22-SEP-1998
DEFINITION HS_3095_B1_F12_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3095 Col=23 Row=L, DNA sequence.
ACCESSION  AQ123154
VERSION     AQ123154.1  GI:3500320
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

```


TITLE
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
JOURNAL
MEDLINE
99380588
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3095 row: L column: 23
Class: BAC ends
High quality sequence stop: 355.
Location/Qualifiers

FEATURES
source
1..355
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate:3095 Col=23 Row=L"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coil DH10B"

BASE COUNT 145 a 83 c 57 g 70 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AQ123154/rev ..
Align seg 1/1 to reverse of: AQ123154 from: 1 to: 355

113 ValSerAlaLeuGlyGlyIlePro 120
|||||
72 GTTCTCTCTTAGGGGATCCCA 49

seq_name: gb_est2:BI777758

seq_documentation_block:

LOCUS BI777758 356 bp mRNA linear EST 26-SEP-2001
DEFINITION EBR008_SQ001_003_R IGF Barley EBR008 library Hordeum vulgare cDNA
clone EBR008_SQ001_003 5', mRNA sequence.
ACCESSION BI777758
VERSION BI777758.1 GI:15780650
KEYWORDS EST.

SOURCE
ORGANISM

Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 356)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
Ramsey, L., Machray, G., Marshall, D. F. M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk
All sequence has a phred quality score of 20 or over
Seq primer: M13 reverse.
Location/Qualifiers

FEATURES
source
1..356

/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBR008_SQ001_003"
/clone_lib="IGF Barley EBR008 library"
/issue_type="Drought stressed root"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from roots of 3 week old drought stressed barley
plants. Developed as part of the barley transcriptome
resources of BBSRC/SEERAD funded cereal IGF (investigating
Gene Function) project."
BASE COUNT 74 a 115 c 105 g 62 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BI777758/rev ..

Align seg 1/1 to reverse of: BI777758 from: 1 to: 356

61 SerThrSerLeuSerLeuArgSer 68
|||||
95 AGCACCTCTCTCTCTCAGATCT 72

seq_name: gb_est2:BM441268

seq_documentation_block:

LOCUS BM441268 357 bp mRNA linear EST 01-FEB-2002
DEFINITION EBed02_SQ002_006_R IGF Barley EBed02 library Hordeum vulgare cDNA
clone EBed02_SQ002_006 5', mRNA sequence.
ACCESSION BM441268
VERSION BM441268.1 GI:18472043
KEYWORDS EST.

SOURCE
ORGANISM

Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 357)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
Ramsey, L., Machray, G., Marshall, D. F. M. and Waugh, R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unit of Genomics
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: rwaugh@scri.sari.ac.uk
All sequence has a phred quality score of 20 or over
Seq primer: M13 reverse.
Location/Qualifiers

FEATURES
source
1..357
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBed02_SQ002_006"
/clone_lib="IGF Barley EBed02 library"
/issue_type="Endosperm"
/dev_stage="8 days post anthesis"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from endosperm tissue dissected from developing
grains (8 days post anthesis) in glasshouse grown barley

plants. Developed as part of the barley transcriptome
resources of BBSRC/SEERAD funded cereal IGF (Investigating
Gene Function) project."

BASE COUNT 61 a 123 c 112 g 61 t

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x BM441268/rev ..

Align seg 1/1 to reverse of: BM441268 from: 1 to: 357

61 SerTherSerleuSerleuArgSer 68
|||||
131 AGCACCCTCTCTCTCTCAGATCT 108

seq_name: gb_est2:c60468

seq_documentation_block:

LOCUS C60468 360 bp mRNA linear EST 22-SEP-1997
DEFINITION C60468 yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk208a11 5', mRNA sequence.

ACCESSION C60468
VERSION C60468.1 GI:2419173
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 360)
; Rhabditidae; Peloderinae; Caenorhabditis.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae

AUTHORS Kohara,Y., Motobashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano

TITLE 'M., Miyata,A. and Nishigaki,A.
Expression map of the C.elegans genome
Unpublished (1996)
JOURNAL Contact: Yuji Kohara
COMMENT Genome Biology Lab.

National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

FEATURES
Location/Qualifiers

1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"

/clone="yk208a11"
/clone_lib="yuji Kohara unpublished CDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

BASE COUNT 106 a 72 c 78 g 104 t

ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x C60468 ..

Align seg 1/1 to: C60468 from: 1 to: 360

137 GUARGLEUHSARGASNARGIU 144
|||||
97 GAAAGCGTCACAGAAACCGTGAG 120

seq_name: gb_est2:c64952

seq_documentation_block:

LOCUS C64952 360 bp mRNA linear EST 22-SEP-1997
DEFINITION C64952 yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk397d10 5', mRNA sequence.

ACCESSION C64952
VERSION C64952.1 GI:2423657
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 360)
; Rhabditidae; Peloderinae; Caenorhabditis.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae

AUTHORS Kohara,Y., Motobashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
'M., Miyata,A. and Nishigaki,A.
Expression map of the C.elegans genome
Unpublished (1996)
JOURNAL Contact: Yuji Kohara
COMMENT Genome Biology Lab.

National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

FEATURES
Location/Qualifiers

1..360
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"

/clone="yk397d10"
/clone_lib="yuji Kohara unpublished CDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

BASE COUNT 105 a 71 c 77 g 107 t

ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x C64952 ..

Align seg 1/1 to: C64952 from: 1 to: 360

137 GUARGLEUHSARGASNARGIU 144
|||||
104 GAAAGCGTCACAGAAACCGTGAG 127

seq_name: gb_est2:D69091

seq_documentation_block:

LOCUS D69091 360 bp mRNA linear EST 07-DEC-1995
DEFINITION CELK059GZF yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk59g12 5', mRNA sequence.

ACCESSION D69091
VERSION D69091.1 GI:1104744
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 360)
; Rhabditidae; Peloderinae; Caenorhabditis.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae

AUTHORS Kohara,Y., Mitsuaki,H., Nishigaki,A., Motobashi,T., Sugimoto,A. and
Tabara,H.
Toward an expression map of the C.elegans genome
Unpublished (1994)
JOURNAL Contact: Yuji Kohara
COMMENT

Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES

source

1. 360
Location/Qualifiers
/organism="Caenorhabditis elegans"
/strain="C41489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="YK59g12"
/clone_lib="Yuji Kohara unpublished cDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT 99 a 71 c 81 g 97 t 12 others
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x D69091 ..

Align seg 1/1 to: D69091 from: 1 to: 360

137 GUUAGTGUHISARGASARGCII 144
|||||
230 GAAAGGCTGCACAGAAACCTGAG 253

seq_name: gb_est1:AW740520

seq_documentation_block:

LOCUS AW740520 363 bp mRNA linear EST 27-APR-2000
DEFINITION BR110477 Blomphalaria glabrata (BS-90)-unexposed lambda zap library
Blomphalaria glabrata cDNA clone RBG1G25TR, mRNA sequence.
ACCESSION AW740520
VERSION AW740520.1 GI:7651613

KEYWORDS

EST.
SOURCE bloodfluke planorb.

ORGANISM

Blomphalaria glabrata
Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora
; Planorbidae; Blomphalaria.

1 (bases 1 to 363)

Raghavan, N., Miller, A., Gardner, M., Kerlavage, A.R., Fitzgerald, P.C.,
Lewis, F.A. and Knight, M.

Genes expressed by the hemocytes of Blomphalaria glabrata before
and after exposure to miracidia
Unpublished (2000)

JOURNAL

Contact: Raghavan N

Biomedical Research Institute
12111 Parklawn Dr., Rockville, MD 20852, USA
Tel: 301-881-3300 ext.128
Fax: 301-770-4756

Email: nkrhelix.nih.gov, snailsrule@aol.com.

Location/Qualifiers

FEATURES

source

1. 363
/organism="Blomphalaria glabrata"
/strain="BS-90"
/db_xref="taxon:6526"
/clone="RBG1G25TR"
/clone_lib="Blomphalaria glabrata (BS-90)-unexposed lambda
zap library"
/sex="hermaphrodite"
/cell_type="Hemocyte"
/lab_host="Laboratory host"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Total RNA was isolated from the hemocytes of
unexposed Blomphalaria glabrata (BS-90) snails and first

strand cDNA synthesized using an oligo-dT primer-linker
(XhoI). Second strand synthesis was followed by the
ligation of EcoRI adaptors. Following digestion with XhoI,
the completed, directional cDNA was cloned into Uni-ZAP
XR phagemid vector by Stratagene.

BASE COUNT 109 a 77 c 71 g 106 t
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AW740520/rev ..

Align seg 1/1 to reverse of: AW740520 from: 1 to: 363

61 SerThSerLeuSerLeuArgSer 68

|||||
186 AGTACTTCACTCTCTCTGAGGAGT 163

seq_name: gb_gss:AQ099964

seq_documentation_block:

LOCUS AQ099964 369 bp DNA linear GSS 27-AUG-1998
DEFINITION HS_3054_B1_B07_MF CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone plate=3054 Col=13 Row=D, DNA sequence.
ACCESSION AQ099964
VERSION AQ099964.1 GI:3470993
KEYWORDS GSS.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 369)

Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, V., Young, V., Zhao, S., Adams, M.D. and
Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

JOURNAL

MEDLINE
99380589

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3054 Row: D Column: 13

Class: BAC ends

High quality sequence stop: 369.

Location/Qualifiers

FEATURES

source

1. 369
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3054 Col=13 Row=D"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelosAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 78 a 100 c 64 g 127 t
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x A0099964 ..
Align seg 1/1 to: A0099964 from: 1 to: 369
64 LeuSerLeuArgSerAlaHisLeu 71
|||||
167 CTTTCCTGAGGCTCTCACCCTC 190
seq_name: gb_gss:A0544348
seq_documentation_block:
LOCUS A0544348 370 bp DNA linear GSS 28-MAY-1999
DEFINITION CITBI-EI-2651G15.TR CITBI-EI Homo sapiens genomic clone 2651G15,
DNA sequence.
ACCESSION A0544348
VERSION A0544348.1 GI:4903423
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 370)
Zhuo,S., Adams,M.D., Niernan,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Other.GSS: CITBI-EI-2651G15.TF
Contact: Shaying Zhuo, William Niernan, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@igr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source location/Qualifiers
1..370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2651G15"
/clone_lib="CITBI-EI"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11, Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT 79 a 65 c 80 g 146 t
ORIGIN
alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x A0544348 ..
Align seg 1/1 to: A0544348 from: 1 to: 370
113 ValSerAlaLeuGlyIlePro 120
|||||
305 GTATCTGCTTTAGGAGGCAATCCCA 328
seq_name: gb_est2:H78351
seq_documentation_block:
LOCUS H78351 372 bp mRNA linear EST 09-NOV-1995

DEFINITION yu79f12.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:240047 5', mRNA sequence.
ACCESSION H78351
VERSION H78351.1 GI:1056440
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 372)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kuwaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Treaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 692
High quality sequence stops: 236
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 692 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 236.
FEATURES
source location/Qualifiers
1..372
/organism="Homo sapiens"
/db_xref="GDB:3789000"
/db_xref="taxon:9606"
/clone="IMAGE:240047"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pTZ19 (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15', AACTGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 76 a 96 c 93 g 103 t 4 others
ORIGIN
alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-528-682-1 x H78351 ..
Align seg 1/1 to: H78351 from: 1 to: 372
113 ValSerAlaLeuGlyIlePro 120
|||||
7 GTCTCTGCAATTAGGGGCAATCCCA 30
seq_name: gb_gss:A2081963
seq_documentation_block:
LOCUS A2081963 372 bp DNA linear GSS 08-JAN-2001
DEFINITION UP-566-2E.T7 RPL11 Human Male BAC Library Homo sapiens genomic
clone 566-2E, DNA sequence.

ACCESSION AZ081963
 VERSION AZ081963.1 GI:7709217
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 372)
 AUTHORS Cheung,V.G., Daitymple,H.L., Narasimhan,S., Watts,J., Schuler,G., Raep,A.K., Morley,M. and Bruzel,A.
 TITLE A resource of mapped human bacterial artificial chromosome clones
 JOURNAL Genome Res. 9 (10), 989-993 (1999)
 MEDLINE 99455100
 COMMENT Contact: Arcaro MA, Morley M, Burdick J, Cheung VC
 Department of Pediatrics
 University of Pennsylvania
 3516 Civic Center Blvd, ARC 516, Philadelphia, PA 19104, USA
 Tel: 215 590 2664
 Fax: 215 590 3709
 Email: mlennox@mail.med.upenn.edu
 Plate: 566 Row: E Column: 2
 Seq primer: 17
 Class: BAC ends.

FEATURES
 source Location/Qualifiers
 1..372
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1b="RPC111 Human Male BAC Library"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; RPC111 Human Male BAC Library"
 BASE COUNT 90 a 69 c 46 g 167 t
 ORIGIN

Alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x AZ081963/rev ..

Align seg 1/1 to reverse of: AZ081963 from: 1 to: 372

217 LysValLysArgGlnIlePheSer 224
 ||||||||||||||||||||
 311 AAGCTAAAGACAAATATTTTCT 288

seq_name: gb_gss:BH462450

seq_documentation_block:
 LOCUS BH462450 381 bp DNA linear GSS 13-DEC-2001
 DEFINITION BOKH235TF BOKH Brassica oleracea genomic clone BOKH235, DNA
 sequence.
 ACCESSION BH462450
 VERSION BH462450.1 GI:17654289
 KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 381)
 REFERENCE Town,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished (2001)
 COMMENT Other_GSSs: BOKH235TR
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523

Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: 17
 Class: sheared ends.

FEATURES
 source Location/Qualifiers
 1..381
 /organism="Brassica oleracea"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BOKH235"
 /clone_1b="BOKH"
 /note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"
 BASE COUNT 135 a 73 c 82 g 91 t
 ORIGIN

Alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-528-682-1 x BH462450/rev ..

Align seg 1/1 to reverse of: BH462450 from: 1 to: 381

60 ValSerThrSerLeuSerLeuArg 67
 ||||||||||||||||||||
 82 GTATCACTTCTCTCAGCCGTCAGC 59

seq_name: gb_est2:BE521751

seq_documentation_block:
 LOCUS BE521751 382 bp mRNA linear EST 19-MAR-2001
 DEFINITION M21C8STM Arabidopsis developing seed Arabidopsis thaliana cDNA
 clone M21C8 5', mRNA sequence.
 ACCESSION BE521751
 VERSION BE521751.1 GI:9779729
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 382)
 REFERENCE White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
 TITLE A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil
 JOURNAL Plant Physiol. 124 (4), 1582-1594 (2000)
 MEDLINE 20567808
 COMMENT Contact: Benning, C
 Dept. of Biochemistry & Molecular Biology
 Michigan State University
 224 Biochemistry, Michigan State University, East Lansing, MI 48824
 , USA
 Tel: 517 355 1609
 Fax: 517 353 9334
 Email: benning@msu.edu
 Michigan State University DNA Sequencing Facility Arabidopsis Biological Resource Center, The Ohio State University, 309 Botany & Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX: 6142920603 TEL: 6142929371.

FEATURES
 source Location/Qualifiers
 1..382
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="M21C8"
 /clone_1b="Arabidopsis developing seed"
 /tissue_type="seed"

alignment_block:
US-09-528-682-1 x W78167/rev ..

Align seg 1/1 to reverse of: W78167 from: 1 to: 393

220 ArgGlnIlePheSerAspTyrGln 227
|||||
69 AGACAATCTTTCTGATTATCA 46

seq_name: gb_gss:AQ114866

seq_documentation_block:
LOCUS AQ114866 395 bp DNA linear GSS 29-AUG-1998
DEFINITION CIT-HSP-2375K8.FR CIT-HSP Homo sapiens genomic clone 2375K8, DNA
sequence.
ACCESSION AQ114866 GI:3490987
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 395)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Title Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Other GSSs: CIT-HSP-2375K8.TF
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source location/Qualifiers
1..395
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2375K8"
/clone_1lb="CIT-HSP"
/sex="male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 102 a 92 c 93 g 108 t

ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AQ114866/rev ..

Align seg 1/1 to reverse of: AQ114866 from: 1 to: 395

113 ValSerAlaLeuGlyGlyIlePro 120
|||||
367 GTATCTGCTTTAGCGGCATCCCA 344

seq_name: gb_gss:AQ120888

seq_documentation_block:
LOCUS AQ120888 396 bp DNA linear GSS 22-SEP-1998
DEFINITION HS-3073.B1.C01.MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-3073 Col-1 Row-F, DNA sequence.
ACCESSION AQ120888 GI:3498054
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 396)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Title Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
JOURNAL MEDLINE
99380589
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3073 Row: F Column: 1
Class: BAC ends
High quality sequence stop: 396.
Location/Qualifiers
1..396
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-3073 Col-1 Row-F"
/clone_1lb="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones In
E-Coli DH10B"

BASE COUNT 145 a 93 c 74 g 80 t 4 others

ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AQ120888/rev ..

Align seg 1/1 to reverse of: AQ120888 from: 1 to: 396

113 ValSerAlaLeuGlyGlyIlePro 120
|||||
100 GTATCTGCTTTAGCGGCATCCCA 77

seq_name: gb_est2:BF908074

seq_documentation_block:
LOCUS BF908074 398 bp mRNA linear EST 18-JAN-2001
DEFINITION RC1-UT0083-091000-013-f01 UT0083 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF908074
VERSION BF908074.1 GI:12299532
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 398)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,

TITLE
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

JOURNAL MEDLINE
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC1&t2=RC1-UT0083-091000-013-f01&t3=2000-10-09&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 397.

FEATURES
source
1..398
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0083"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
93 a 93 c 70 g 142 t

ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x BF908074/rev ..

Align seg 1/1 to reverse of: BF908074 from: 1 to: 398

61 SerTherSerLeuSerLeuArgSer 68
|||||
31 TCCACCTCCCTCAGCTGAGATCA 8

seq_name: gb_est1:AM587771

seq_documentation_block:
LOCUS AM587771 400 bp mRNA linear EST 22-MAR-2000
DEFINITION ST66F07 Pine Triplex shoot tip library Pinus taeda cDNA clone
ACCESSION AM587771
VERSION AM587771.1 GI:7274798
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 400)
REFERENCE
1 Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
The Pine Gene Discovery Project
Unpublished (1999)
CONTACT: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh

TITLE
, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhetten@uncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing primer.
Location/Qualifiers

FEATURES
source
1..400
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="ST66F07"
/clone_lib="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site_1: SfiI (A); Site_2: SfiI (B); Shoot tips (approx. 2 cm from apex) were collected during the spring, frozen and used for mRNA isolation. The SMART-PCR method (Clontech) was used to prepare a library from 1 ug total RNA, using the Lambda Triplex vector. Plasmid subclones in pTriplex were recovered by cre-lox excision in E. coli strain BM25.8 and sequenced from the 5' end."

BASE COUNT
81 a 98 c 110 g 97 t 14 others

ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AM587771 ..

Align seg 1/1 to: AM587771 from: 1 to: 400

14 AspGluIlelySArgSerGlyGly 21
|||||
244 GATGAGATCAAGCGCGGCGGG 267

seq_name: gb_est1:AM902939

seq_documentation_block:
LOCUS AM902939 403 bp mRNA linear EST 24-MAY-2000
DEFINITION QV3-NN1025-100500-183-b01 NN1025 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM902939
VERSION AM902939.1 GI:8067144
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 403)
REFERENCE
1 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M.J., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

JOURNAL MEDLINE
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=st2-QV3-NN1025-100500-183-b01&t3=2000-05-10&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 18
High quality sequence stop: 403.
Location/Qualifiers

FEATURES

source

1. 403
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="NN1025"
/dev_stage="Adult"
/note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 138 a 95 c 90 g 80 t

ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x AW902939/rev ..

Align seg 1/1 to reverse of: AW902939 from: 1 to: 403

113 ValSerAlaLeuGlyGlyIlePro 120
|||||
74 GTCTCTGCTTAGGGGCAATCCCA 51

seq_name: gb_est1:A1350210

seq_documentation_block:

LOCUS A1350210 408 bp mRNA linear EST 01-FEB-1999
DEFINITION Q038A10.X1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1910298 3',
mRNA sequence.

ACCESSION A1350210 GI:4087416

VERSION EST.

KEYWORDS human.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 566 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 379.
Location/Qualifiers

FEATURES

source

1. 408
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="NCI_CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"

/note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker. 1st strand cDNA was prepared from neuroendocrine lung carcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 163 a 62 c 61 g 122 t

ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x A1350210 ..

Align seg 1/1 to: A1350210 from: 1 to: 408

63 SerLeuSerLeuArgSerAlaHis 70
|||||
227 TCATTATCTTACGACAGTGCACAC 250

seq_name: gb_est2:BF777429

seq_documentation_block:

LOCUS BF777429 410 bp mRNA linear EST 12-JAN-2001
DEFINITION NXSL_071_E05_F NXSL (NsF Xylem Side wood Inclined) Pinus taeda cDNA
clone NXSL_071_E05 5', mRNA sequence.

ACCESSION BF777429

VERSION BF777429.1 GI:12125329

KEYWORDS EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Johnson, Arthur

North Carolina State University

Tel: 919 515 7800

Fax: 919 515 7801

Email: ajohnson@unity.ncsu.edu

Seq primer: T3

Location/Qualifiers

1. 410

/organism="Pinus taeda"

/strain="Coastal plain loblolly pine from North Carolina"

/db_xref="taxon:3352"

/clone_1lb="NXSL_071_E05"

/tissue_type="Xylem"

/cell_type="Side"

/dev_stage="Juvenile"

/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI; The library is from early (spring) wood, taken from three six-year old trees (three different genotypes), in the juvenile phase. These trees were induced to form side wood by bending to a 45 degree angle and tying them to the ground. Differentiating xylem was harvested from the sides of the inclined stems, and a mixture of all three genotypes was used for the library. oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI Bluescript SK vector arms. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGCGACGAG'."

BASE COUNT

81 a 102 c 107 g 105 t 15 others

BASE COUNT 104 a 90 c 74 g 149 t
ORIGIN

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x AZ768672 ..

Align seg 1/1 to: AZ768672 from: 1 to: 417

59 TyrValSerThrSerLeuSerLeu 66
|||||
393 TACGTATCCACATCTCTCTCTCTC 416

[illegible]

/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-255-456-35705	+ 8.00	137.66	323.22	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-255-457-11066	+ 8.00	137.66	323.28	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-255-456-36043	+ 8.00	137.66	324.26	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-378-1981	+ 8.00	137.59	326.21	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-26906	+ 8.00	137.59	326.21	8.00
/csgn2.6/p.todata/2/pna/US094.COMB.seq:US-09-411-999-12231	+ 8.00	137.54	328.17	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-10514	+ 8.00	137.52	329.14	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-255-456-34883	+ 8.00	137.49	330.12	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-09-605-701-1597	+ 8.00	137.47	331.09	8.00
/csgn2.6/p.todata/2/pna/US06014.COMB.seq:US-60-144-084-30439	+ 8.00	137.47	331.09	8.00
/csgn2.6/p.todata/2/pna/US0937B.COMB.seq:US-60-253-456-33482	+ 8.00	137.47	331.09	8.00
/csgn2.6/p.todata/2/pna/US0937B.COMB.seq:US-09-724-750-33862	+ 8.00	137.42	333.04	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-456-33460	+ 8.00	137.42	333.04	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-255-457-48196	+ 8.00	137.42	333.04	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-48196	+ 8.00	137.40	334.02	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-26400	+ 8.00	137.40	334.02	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-46637	+ 8.00	137.40	334.02	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-456-31414	+ 8.00	137.27	339.86	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-42682	+ 8.00	137.24	340.83	8.00
/csgn2.6/p.todata/2/pna/US0964.COMB.seq:US-09-605-775-4367	+ 8.00	137.22	341.80	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-255-378-33678	+ 8.00	137.22	341.80	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-09-605-456-34979	+ 8.00	137.20	342.78	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-755-4446	+ 8.00	137.20	342.78	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-09-523-457-44231	+ 8.00	137.16	343.75	8.00
/csgn2.6/p.todata/2/pna/US06014.COMB.seq:US-60-140-769-33059	+ 8.00	137.16	344.72	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-47653	+ 8.00	137.11	346.66	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-255-457-48067	+ 8.00	137.07	349.58	8.00
/csgn2.6/p.todata/2/pna/US0964.COMB.seq:US-09-605-702-20066	+ 8.00	137.02	350.55	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-701-20770	+ 8.00	137.00	351.52	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-47040	+ 8.00	137.00	351.52	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-47040	+ 8.00	136.98	352.49	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-456-33077	+ 8.00	136.90	356.36	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-09-605-702-19647	+ 8.00	136.88	357.33	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-60-255-378-12384	+ 8.00	136.88	357.33	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-10263	+ 8.00	136.78	361.20	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-701-20297	+ 8.00	136.73	364.11	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-25721	+ 8.00	136.73	364.11	8.00
/csgn2.6/p.todata/2/pna/US0932.COMB.seq:US-09-770-848B-107743	+ 8.00	136.71	365.07	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-255-457-25382	+ 8.00	136.69	366.04	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-25382	+ 8.00	136.67	367.01	8.00
/csgn2.6/p.todata/2/pna/US0937B.COMB.seq:US-09-724-730-6363	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US06017.COMB.seq:US-09-717-431-6363	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-25637	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-25637	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US06025.COMB.seq:US-60-253-457-25637	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-605-702-19487	+ 8.00	136.63	368.94	8.00
/csgn2.6/p.todata/2/pna/US0936A.COMB.seq:US-09-				


```

/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-456-16766 + 8.00 136.21 389.17
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-866-7756 + 8.00 136.17 391.10
/cgn2_6/ptodata/2/pna/US0917_COMB.seq:US-09-724-866-7756 + 8.00 136.17 391.10
/cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-60-171-432-7736 + 8.00 136.17 391.10
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-606-755-15023 + 8.00 136.13 393.02
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-606-755-15023 + 8.00 136.11 393.98
/cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-654-617-33319 - 8.00 136.11 393.98
/cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-684-016-33319 - 8.00 136.11 393.98
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-26284 + 8.00 136.10 394.94
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-702-19448 + 8.00 136.08 395.90
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-27106 + 8.00 136.06 396.86
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-701-9421 + 8.00 136.06 396.86
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-750-10210 + 8.00 136.04 397.82
/cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-60-171-431-10210 + 8.00 136.04 397.82
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-9616 + 8.00 136.02 398.78
/cgn2_6/ptodata/2/pna/PCNUS_COMB.seq:PCN-US01-01333-606 + 8.00 136.02 398.78
/cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-764-860-606 + 8.00 136.02 398.78
/cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-764-860-606 + 8.00 136.02 398.78
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-074-095-606 + 8.00 136.02 398.78
/cgn2_6/ptodata/2/pna/US100_COMB.seq:US-10-091-548-78 + 8.00 136.02 398.78
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-26276 + 8.00 136.00 399.73
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-46900 + 8.00 135.96 401.65
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-274-861-798 + 8.00 135.95 402.61
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-915-738-798 + 8.00 135.95 402.61
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-750-3352 + 8.00 135.91 404.53
/cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-60-171-431-3352 + 8.00 135.91 404.53
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-701-9282 + 8.00 135.89 405.49
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-750-18758 + 8.00 135.87 406.44
/cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-60-171-431-18758 + 8.00 135.87 406.44
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-25780 + 8.00 135.73 414.10
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-909-629-23289 + 8.00 135.69 416.01
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-909-629-23289 + 8.00 135.65 416.01
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-701-1678 + 8.00 135.65 417.92
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-950-083-10248 - 8.00 135.58 422.64
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-26363 + 8.00 135.57 422.64
/cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-865-555-16625 + 8.00 135.55 423.65
/cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-565-306-71809 + 8.00 135.53 424.60
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-411-989-14963 + 8.00 135.50 426.51
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-701-91358 + 8.00 135.50 426.51
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-701-15370 + 8.00 135.50 426.51
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-606-755-15017 + 8.00 135.48 427.46
/cgn2_6/ptodata/2/pna/US092_COMB.seq:US-09-289-768-3050 - 8.00 135.46 428.42
/cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-939-397-3050 - 8.00 135.46 428.42
/cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-667-188A-4042 + 8.00 135.46 428.42
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-866-9741 + 8.00 135.41 431.28
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-866-9741 + 8.00 135.41 431.28
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-866-9741 + 8.00 135.41 431.28
/cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-60-155-006-4042 + 8.00 135.41 431.28
/cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-60-171-432-9741 + 8.00 135.41 431.28
/cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-331-534-10390 - 8.00 135.34 435.08
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-606-755-15461 - 8.00 135.34 435.08
/cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-465-877-8124 + 8.00 135.31 436.99
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-906-555-8124 + 8.00 135.31 436.99
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-606-755-15192 + 8.00 135.29 437.89
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-25068 + 8.00 135.27 438.84
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-27170 + 8.00 135.17 444.59
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-750-6137 + 8.00 135.11 448.39
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-750-10136 + 8.00 135.11 448.39
/cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-60-171-431-6137 + 8.00 135.11 448.39
/cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-60-171-431-10136 + 8.00 135.11 448.39
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-902-540-6102 + 8.00 135.09 449.34
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-702-19824 + 8.00 134.97 455.98
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-701-9701 + 8.00 134.96 456.93
/cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-605-701-15649 + 8.00 134.94 457.88
/cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-912-292-252198 + 8.00 134.91 459.77
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-456-31737 + 8.00 134.81 465.45
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-750-3100 + 8.00 134.74 470.18
/cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-724-750-12188 + 8.00 134.74 470.18
/cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-60-171-431-12188 + 8.00 134.74 470.18
/cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-60-253-457-14198 + 8.00 134.74 470.18
/cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-60-168-611-23321 - 8.00 134.61 477.74

```


[illegible]

/cgn2_6/p.todata/2/pna/US6022.COMB	.seq	.seq	US-60-229-555-62	+	8.00	104.25	2.3e+04	3221	/cgn2_6/p.todata/2/pna/US099B.COMB	.seq	.seq	US-09-950-335A-7	-	7.00	132.85	598.91
/cgn2_6/p.todata/2/pna/US6020.COMB	.seq	.seq	US-60-207-316-4	+	8.00	104.22	2.4e+04	3235	/cgn2_6/p.todata/2/pna/PCrUS.COMB	.seq	.seq	PCr-US99-30747-66	+	7.00	132.18	652.39
/cgn2_6/p.todata/2/pna/US6020.COMB	.seq	.seq	US-60-200-382-25	+	8.00	104.13	2.4e+04	3273	/cgn2_6/p.todata/2/pna/PCrUS.COMB	.seq	.seq	PCr-US99-30747-66	+	7.00	132.18	652.39
/cgn2_6/p.todata/2/pna/US6020.COMB	.seq	.seq	US-60-200-382-26	+	8.00	104.13	2.4e+04	3273	/cgn2_6/p.todata/2/pna/US093.COMB	.seq	.seq	US-09-313-221A-88	+	7.00	132.18	652.39
/cgn2_6/p.todata/2/pna/US6020.COMB	.seq	.seq	US-60-207-315-61	+	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US094.COMB	.seq	.seq	US-09-470-124-66	+	7.00	132.18	652.39
/cgn2_6/p.todata/2/pna/US6020.COMB	.seq	.seq	US-60-208-020-8	+	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US094.COMB	.seq	.seq	US-09-470-124-66	+	7.00	132.18	652.39
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-209-004-9	+	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/PCrUS.COMB	.seq	.seq	PCr-US01-31138-8	-	7.00	130.61	797.59
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-97	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-112A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-98	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.seq	US-60-213-178-99	-	8.00	104.12	2.4e+04	3276	/cgn2_6/p.todata/2/pna/US099A.COMB	.seq	.seq	US-09-908-975-113A2	-	7.00	130.12	849.79
/cgn2_6/p.todata/2/pna/US6021.COMB	.seq	.														


```

4337 TTCTCAGATATATAGATGGATGCTGTTATTTTGGTGTGATGATGAC 4386
138 ryleuhlsarqasnaargluTyTArqAsprgrtyTyTArqAsnleuasn 154
4387 GATTACATCGTAAcAGGGAATTATAGAGACCcGGATTAACAGAAATCTAAT 4436
155 lIeAlaPrOAlaGluaSpGlYTyTArqleuAlaGlYpheProPAspHl 171
4437 ATAGCTCCGGAGAGAGATGGTTACAGATTAGCAGGTTTCCACCGGAWCA 4486
171 sGlnAlaTPrArqGluaGluprOrPrIeHlshIsAlaPrOgluGlYysg 188
4487 CCAGGTTGGAGAAAGAACCCtGGATTCATGCAACCCACCAAGGTTGTG 4536
188 lYAsnSerSerArqThrlIethrGlYAsPrHrcYsaSngluGlutHrGlN 204
4537 GAAATTCATCCAGAACCAATATACAGGCGATCTGTATATGAGAGACCCAG 4586
205 AsnleuSerThrlIethrleuauG 212
4587 AATCTGAGCAATATATCTCAGG 4610

seq_name: /cgm2_6/ptodata/2/pna/US097B_-COMB.seq:US-09-724-315-6
seq_documentation_block:
: Sequence 6, Application US/09724315
: GENERAL INFORMATION:
: APPLICANT: HAYNES, Joel R.
: APPLICANT: ARRINGTON, Joshua
: TITLE OF INVENTION: NOCLEIC ACID ADJUVANTS
: FILE REFERENCE: App41

```

```

; NUMBER OF SEQ ID NOS: 22
;
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 6
;
; LENGTH: 5408

```

```

; ORGANISM: plasmid pPJV2007
US-09-724-315-6

```

Quality:	186.00	Length:	186
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

```
alignment_block:
US-09-528-682-1 x US-09-724-315-6 ..
```

27 HisAsnGluTyrPheaspargGlyThrGlnMetAsnIleAsnLeuTyrAs 43

43 pHisAlaArgGlyThrGlnThrGlyPheValArgTyrAspSpGlyTyr 60
4053 CAATAAGACATCTCGATAGAGGAACTCAATGAAATTAATCTTAAGA 4102

60 a) Ser Thr Ser Leu Ser Leu Arg Ser Ala His Leu Ala Gly Ile 76
4103 TCACGCGAGAGGAACACCAACCGGCTTTGTCA GATATGATGACCGGATATG 4152

77 leuSerGIyTyrSerThrTyrTyrIleTyrValIlealathrAlaProAs 93
4153 TTTCCACCTTCCTTAGTTTGAGAAAGTGCTCACCCTTAGCAGGACAGTCTATA 42022

4203 TTATCAGGATATTCACCTTACTATATATATGTATTAGCGACAGCACCAAA 4252
93 nmetpheasnValasnaSprValleuGlyValTyrSerProHisProTyrG 110

4253 TATGTTAATGTTAATGATGTAATAGGCGTATACAGCCCTCACCACATATG 4302


```

110 lucIngluValSerAlaLeuGlyGlyIleProTyrSerGlnIleTyrGly 126
|||||
4303 AACAGAGAGGTTTCTGCTTAGCGTAGGATACCATATTTCTCAGATATATAGA 4352
127 TyrTyrArgValAsnPhgGlyValIleAspGluArgLeuHisArgAsnAr 143
|||||
4353 TGGTATCGTGTATATTTGGTGTGATGTAGAGATTTACATGCTAGACAG 4402
143 gGluTyrArgAspArgTyrTyrArgAsnLeuAsnIleAlaProAlaGlu 160
|||||
4403 GGAATATAGAGACCGGTATTACAGAAATTCGAATATAGCTCCGACAGAG 4452
160 spGlyTyrArgLeuAlaGlyPheProProAspHisGlnAlaTTrpArgGlu 176
|||||
4453 ATGGTTACACATATAGCAGGTTTCCACCGCATCACCAAGCTTGAGAGAA 4502
177 GluProTTrpIleHisAlaProGlnGlyCysGlyAsnSerSerArgTh 193
|||||
4503 GAACCCGTGATCATCATGCACCAAGAGTTGTGAAATTCATCAAGAAC 4552
193 rIleThGlyAspThrCysAsnGluGluThrGlnAsnLeuSerThrIleT 210
|||||
4553 AATTACAGGTACTGTGTATATAGAGAGACCAAGATTCAGACCAATAT 4602
210 yLeuArg 212
|||||
4603 ATCTCAG 4610

seq_name: /cgn2_6/ptodata/2/pna/PCtUS.COMB.seq: PCT-US99-30747-3

seq_documentation_block:
; Sequence 3, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(782)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-3

alignment_scores:
Quality: 184.00 Length: 184
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-3 ..
Align seg 1/1 to: PCT-US99-30747-3 from: 1 to: 782

5 LeuTyrArgAlaAspSerArgProProAspGluIleLeuArgSerGlyG 21
|||||
72 CTCTATAGGCGTACTGTAGCCGCCAGATGATCAAGAGGCTGAGG 121
21 yLeuMetProArgGlyHisAsnGluTyrPheAspArgGlyThGlnMet 38
|||||
122 TCTCATGCCAAGGGAGCACAAATGATGACTTTGATAGGGGAACCAATGA 171

```

```

38 snIleAsnLeuTyrAspHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
172 ACATCAACCTCTATGACCATGCTTAGGGGAACCCAAATGCTGCTTAGG 221
55 TyrAspGlyTyrValSerThrSerLeuSerLeuArgSerAlaHisLe 71
|||||
222 TATGATGATGGAATATGTGTCACCTCCCTTAGCTTGAGAGCTGCTCACT 271
71 uAlaGlyGlnSerIleLeuSerGlyTyrSerThrTyrTyrIleTyrVal 88
|||||
272 GCGCAGGCAATCCATCCCTCAGATATCTCCACCTTACTATCATATGTA 321
88 leaIarTrpAlaProAsnMetPheAsnValAsnAspValIleGlyValTyr 104
|||||
322 TTGCTTACAGACCAAAATGTTCAATGTGAATGATGTGTGGAGTAT 371
105 SerProHisProTyrGluGlnGluValSerAlaLeuGlyIleProTyr 121
|||||
372 AGCCCTCACCATATGAGCAAGAGGTGTGCTTGGGTGGAATCCCAT 421
121 rSerGlnIleTyrGlyTTrpTyrArgValAsnPhgGlyValIleAspGlu 138
|||||
422 CTCCCAATCTATGATGATGATAGGTTGAACCTTGTGTGATGATGAGA 471
138 rGleuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsnLeuAsn 154
|||||
472 GCGTCCATAGGAATAGGAGATATAGGACAGTACTATAGAACCTTCAAC 521
155 IleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProProAspH 171
|||||
522 ATACCTCCAGCAGAGAGATGTTATAGTTGGCAGGTTCCACACAGACA 571
171 sGlnAlaTTrpArgGluGluProTTrpIleHisAlaProGlnIleTyr 188
|||||
572 CCAAGCCTGGAGGAGAGCCCTGTGATCCACATGACACCAAGTTGTG 621
188 y 188
||
622 GA 623

seq_name: /cgn2_6/ptodata/2/pna/US094.COMB.seq: US-09-470-124-3

seq_documentation_block:
; Sequence 3, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(782)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-3

alignment_scores:
Quality: 184.00 Length: 184
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```


alignment_block:

US-09-528-682-1 x US-09-470-124-3

Align seg 1/1 to: US-09-470-124-3 from: 1 to: 782

```

5 LeuTyraArgAlaaspSerArgProProaspGluLeuLysArgSerGly 21
|||||
72 CTCTATAGGCTGACTAGAGCCGCCAGATGATCAAGAGCTGTGAGG 121
21 yLeuMetProArgGlyHisAsnGlyuTyrrPheaspArgGlyThrClnMet 38
|||||
122 TCTCATGCCAAGGGGACAAATGAGTACTTGTATGAGGGAACCCAAATGA 171
38 snlLeuAsnLeuTyraSPHisAlaArgGlyThrGlnThrGlyPheValArg 54
|||||
172 ACATCAACCTCTATGACCATGCTAGGGGAACCCAACTGGCTTGTGAGG 221
55 TyraSPaspGlyTyrrValSerThrSerLeuSerLeuArgSerAlaHisLe 71
|||||
222 TATGATGATGATATGTGTCACCTCCCTTAGCTTGAAGTCTGCTCACTT 271
71 uAlaGlyGlnSerIleLeuSerGlyTyrrSerThrTyrrIleTyrrVal 88
|||||
272 GGCAGGACAATCCATCCCTCAGGATGATCCACTACTACATCTATGTA 321
88 leaIaThrAlaProAsnMetPheAsnValAsnAspValIleuGlyValTyrr 104
|||||
322 TTGCTACAGCACAACATGTTCAATGTAATGATGTGTGGAGTGTAT 371
105 SerProHisProTyrrGluGlnGluValSerAlaLeuGlyIleProTy 121
|||||
372 AGCCCTCAACCATATGAGCAAGAGGTCTGCTTGGTGGTAATCCATA 421
121 rSerGlnIleTyrrGlyTyrrArgValAsnPheGlyValIleaspGlu 138
|||||
422 CTCCCAATCTATGATGATGATGAGGTAAGTGTGTTGATGATGAGA 471
138 rGleuHisArgAsnArgIleuTyrrArgAspArgTyrrArgAsnLeuAsn 154
|||||
472 GGCTCCATAGGATAGGAGATAGGAGGACAGTACTATAGAACCTCAAC 521
155 lIleAlaProAlaGluaspGlyTyrrArgLeuAlaGlyPheProProAsp 171
|||||
522 ATACTCCAGCAGAGAGATGTTATAGTGTGACAGTTCACACAGCA 571
171 sGlaIaTrpArgGluGluProTrpIleHisAlaProGlnGlyCysG 188
|||||
572 CCAAGCTGGAGGAGAGACCTGATCCACATGACACACAGAGTGTG 621
188 ly 188
||
622 GA 623

```

seq_name: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-950-335A-5

seq_documentation_block:

```

; Sequence 5, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTION
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-950-335A-5

```

alignment_scores:

Quality: 111.00 Length: 111
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-950-335A-5

Align seg 1/1 to: US-09-950-335A-5 from: 1 to: 723

```

1 AsnGlyAspArgLeuTyrrArgAlaaspSerArgProProaspGluLeu 17
|||||
1 AATGCCACAGATATATACCTGCTGACTAGACCCGCCAGATGAATATAA 50
17 sArgSerGlyGlyLeuMetProArgGlyHisAsnGlyuTyrrPheaspArg 34
|||||
51 ACGTCCGAGGCTTATGCCCAGAGGGCATATGAGTACTTGATGATAGAG 100
34 lYThrGlnMetAsnIleAsnLeuTyraSPHisAlaArgGlyThrGlnThr 50
|||||
101 GAACCTCAATGATATATATCTTATGATCAGCGAGAGACACAAC 150
51 GlyPheValArgTyraSPaspGlyTyrrValSerThrSerLeuSerLeu 67
|||||
151 GCGTTGTGACAGATATGATGAGGATATGTTCCACTTCTCTTGTGTGAG 200
67 gSerAlaHisLeuAlaGlyGlnSerIleLeuSerGlyTyrrSerThrTyrr 84
|||||
201 AAGGTCTACTTACAGCAGCAGCTATATATATCAAGATATCCACTTACT 250
84 yIleTyrrValIleAlaThrAlaProAsnMetPheAsnValAsnAspVal 100
|||||
251 ATATATATGTTATAGCAGCAGACACAATATGTTTATGTTATGATGAT 300
101 LeuGlyValTyrrSerProHisProTyrrGluGln 111
|||||
301 TTAGCGCTATACAGCCCTCACCATATGAAACAG 333

```

seq_name: /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-08-256-003-5

seq_documentation_block:

```

; Sequence 5, Application US/08256003
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera toxin and of the toxin Lt, Their Preparation and
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,003
; FILING DATE: 11-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315,001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) (655-3542

```



```
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..711
US-08-256-003-5

alignment_scores:
    Quality: 86.00      Length: 86
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-256-003-5  ..

Align seg 1/1 to: US-08-256-003-5 from: 1 to: 711

103 ValTyrSerProHisProTyrGluGluGluValSerAlaLeuGlyGlyI 119
|||||
295 GTATACAGCCCTCACCACCATATGACAGAGAGGTTTCGCGTTAGTGGAAT 344
119 eProTyrSerGlnIleTyrGlyTTrpTyrArgValAsnPhcGlyValIleA 136
|||||
345 ACCATATTCTCAGATATATGATGATGATGATGATTTTGGTGTGATTTG 394
136 sPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACGATTACATCGTACAGAGGAATATAGAGACCGGATTATACAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169
|||||
445 CTGATATATAGCTCCGCGACAGAGATGTTACAGATTAGCAGGTTCCACAC 494
169 oAspHisGlnAlaTTPArgGluGluProTPrIleHisAlaProGlnG 186
|||||
495 GGATCACCAGCTTGAGAGAGAAACCCCTGATTCATCATGCACCAAG 544
186 lYcysGly 188
|||||
545 GTTGTGGA 552

seq_name: /cgn2/6/plodata/2/pna/US090_COMB.seq:US-09-044-696-1

seq_documentation_block:
; Sequence 1, Application US/09044696
; GENERAL INFORMATION:
; APPLICANT: BARCHFIELD, GAIL
; APPLICANT: DEL GIUDICE, GIUSEPPE
; APPLICANT: RAPPUOLI, RINO
; TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL
; TITLE OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
; ADDRESS: R440
; STREET: P.O. BOX 8097
; CITY: EMERYVILLE
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/044,696
```

```
; FILING DATE: 18-MAR-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,227
; FILING DATE: 21-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 1393,002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 655-8730
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..708
US-09-044-696-1

alignment_scores:
    Quality: 86.00      Length: 86
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-044-696-1  ..

Align seg 1/1 to: US-09-044-696-1 from: 1 to: 711

103 ValTyrSerProHisProTyrGluGluGluValSerAlaLeuGlyGlyI 119
|||||
295 GTATACAGCCCTCACCACCATATGACAGAGAGGTTTCGCGTTAGTGGAAT 344
119 eProTyrSerGlnIleTyrGlyTTrpTyrArgValAsnPhcGlyValIleA 136
|||||
345 ACCATATTCTCAGATATATGATGATGATGATGATTTTGGTGTGATTTG 394
136 sPGLuArgLeuHisArgAsnArgGluTyrArgAspArgTyrTyrArgAsn 152
|||||
395 ATGACGATTACATCGTACAGAGGAATATAGAGACCGGATTATACAGAAAT 444
153 LeuAsnIleAlaProAlaGluAspGlyTyrArgLeuAlaGlyPheProPr 169
|||||
445 CTGATATATAGCTCCGCGACAGAGATGTTACAGATTAGCAGGTTCCACAC 494
169 oAspHisGlnAlaTTPArgGluGluProTPrIleHisAlaProGlnG 186
|||||
495 GGATCACCAGCTTGAGAGAGAAACCCCTGATTCATCATGCACCAAG 544
186 lYcysGly 188
|||||
545 GTTGTGGA 552

seq_name: /cgn2/6/plodata/2/pna/US099B_COMB.seq:US-09-950-335A-2

seq_documentation_block:
; Sequence 2, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTIO
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
```


LENGTH: 582
TYPE: DNA
ORGANISM: Vibrio cholerae
US-09-950-335A-2

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-950-335A-2 ..

Align seg 1/1 to: US-09-950-335A-2 from: 1 to: 582

29 GUTYRPHASPARGLYTHRGLINMETASNIIEASNLEUTYRASPISAL 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATACACCTTATGATCATGC 134

45 AARGGLYTHRGLNTHRGlyPheValArg 54
|||||
135 AAGAGGAACTCAGACGGGATTTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-08-256-003-7

seq_documentation_block:

Sequence 7, Application US/08256003
GENERAL INFORMATION:
APPLICANT: Domenighini, Mario
APPLICANT: Rappuoli, Rino
TITLE OF INVENTION: Immunogenic Detoxified Mutants of
TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,003
FILING DATE: 11-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0315,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..723
US-08-256-003-7

alignment_scores:

Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-256-003-7 ..

Align seg 1/1 to: US-08-256-003-7 from: 1 to: 723

29 GUTYRPHASPARGLYTHRGLINMETASNIIEASNLEUTYRASPISAL 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATACACCTTATGATCATGC 134

45 AARGGLYTHRGLNTHRGlyPheValArg 54
|||||
135 AAGAGGAACTCAGACGGGATTTGTTAGG 162

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-044-696-3

seq_documentation_block:

Sequence 3, Application US/09044696
GENERAL INFORMATION:
APPLICANT: BARCHFIELD, GAIL
APPLICANT: DEL GUIDICE, GIUSEPPE
TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL
TITLE OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
STREET: P.O. BOX 8097
CITY: EMERYVILLE
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/044,696
FILING DATE: 18-MAR-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,227
FILING DATE: 21-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: HARBIN, ALISA A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 1393,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 655-8730
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..720
NAME/KEY: CDS
LOCATION: 1..720
US-09-044-696-3

alignment_scores:
Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-044-696-3 ..

Align seg 1/1 to: US-09-044-696-3 from: 1 to: 723

```
29 GluTyrPheasparglyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTTATATGATCATGCG 134
|||||
45 aargglyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACTCAAGACGGATTGTGTAGG 162
```

seq_name: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-950-335A-1

seq_documentation_block:

```
; Sequence 1, Application US/09950335A
; GENERAL INFORMATION:
; APPLICANT: HONE, DAVID M.
; TITLE OF INVENTION: GENETICALLY ENGINEERED CO-EXPRESSION DNA VACCINES, CONSTRUCTION
; FILE REFERENCE: 4115-128
; CURRENT APPLICATION NUMBER: US/09/950,335A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Vibrio cholerae
US-09-950-335A-1
```

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-950-335A-1 ..

Align seg 1/1 to: US-09-950-335A-1 from: 1 to: 723

```
29 GluTyrPheasparglyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
85 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTTATATGATCATGCG 134
|||||
45 aargglyThrGlnThrGlyPheValArg 54
|||||
135 AAGAGGAACTCAAGACGGATTGTGTAGG 162
```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-45

seq_documentation_block:

```
; Sequence 45, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
```

LOCATION: (1)..(777)

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-45
```

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-45 ..

Align seg 1/1 to: PCT-US99-30747-45 from: 1 to: 777

```
29 GluTyrPheasparglyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACCGAGGTACTCAAGATGACATCAACCTTTATATGACATGCG 188
|||||
45 aargglyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACCTCAAACTGATTTGTGAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-47

seq_documentation_block:

```
; Sequence 47, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-47
```

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-47 ..

Align seg 1/1 to: PCT-US99-30747-47 from: 1 to: 777

```
29 GluTyrPheasparglyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACCGAGGTACTCAAGATGACATCAACCTTTATATGACATGCG 188
|||||
45 aargglyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACCTCAAACTGATTTGTGAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-49


```
seq_documentation_block:
; Sequence 49, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Transgenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-49
```

```
alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x PCT-US99-30747-49 ..

```
Align seg 1/1 to: PCT-US99-30747-49 from: 1 to: 777
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGCGGCTACTCAGATGACATCAACCTTTATGACCATGC 188
45 aargglythrGlnThrGlyPheValarg 54
|||||
189 AAGGGAAGCTCAAGCTGATTTGTGTAGG 216

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-51

seq_documentation_block:
; Sequence 51, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-51
```

```
alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
```

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x PCT-US99-30747-51 ..

Align seg 1/1 to: PCT-US99-30747-51 from: 1 to: 777

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGCGGCTACTCAGATGACATCAACCTTTATGACCATGC 188
45 aargglythrGlnThrGlyPheValarg 54
|||||
189 AAGGGAAGCTCAAGCTGATTTGTGTAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-07-694-733-1

```
seq_documentation_block:
; Sequence 1, Application US/07694733
; GENERAL INFORMATION:
; APPLICANT: Burnette, W. Neal
; APPLICANT: Kaslow, Harvey R.
; TITLE OF INVENTION: Recombinant DNA-derived
; TITLE OF INVENTION: Cholera Toxin
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: Amgen Center
; STREET: 1640 Dehaven Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in., DS, 1.4 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh OS 7.0.
; SOFTWARE: Microsoft Word Version 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/694,733
; FILING DATE: 19910502
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 777 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double stranded
; TOPOLOGY: circular
; US-07-694-733-1
```

```
alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-07-694-733-1 ..

Align seg 1/1 to: US-07-694-733-1 from: 1 to: 777

```
29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGCGGCTACTCAGATGACATCAACCTTTATGACCATGC 188
45 aargglythrGlnThrGlyPheValarg 54
|||||
189 AAGGGAAGCTCAAGCTGATTTGTGTAGG 216
```

seq_name: /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-08-435-605-1

```
seq_documentation_block:
; Sequence 1, Application US/08435605
```


GENERAL INFORMATION:
APPLICANT: Burnette, W. Neal
APPLICANT: Kaslow, Harvey R.
TITLE OF INVENTION: Recombinant DNA-derived
TITLE OF INVENTION: Cholera toxin
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: Amgen Center
STREET: 1840 Dehavenland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in., DS, 1.4 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh OS 7.0.
SOFTWARE: Microsoft Word Version 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,605
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/694,733
FILING DATE: 02-MAY-1991
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 777 base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: circular
US-08-435-605-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-435-605-1 ..

Align seg 1/1 to: US-08-435-605-1 from: 1 to: 777

29 GluTyrPheaspArgGlyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGAGGCTACTCAATGAAATGCAACCTTTATGTCATGC 188
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAAGACGGGATTGTGAGG 216

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-470-124-45

seq_documentation_block:
Sequence 45, Application US/09470124
GENERAL INFORMATION:
APPLICANT: Mason
APPLICANT: Arntzen
TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
TITLE OF INVENTION: Transgenic Plants
FILE REFERENCE: 4868/84454
CURRENT APPLICATION NUMBER: US/09/470,124
CURRENT FILING DATE: 1999-12-24
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 45
LENGTH: 777
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS

LOCATION: (1)..(777)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
OTHER INFORMATION: cholera toxin gene mutagenized to optimize
OTHER INFORMATION: expression in plants.
US-09-470-124-45

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-470-124-45 ..

Align seg 1/1 to: US-09-470-124-45 from: 1 to: 777

29 GluTyrPheaspArgGlyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGAGGCTACTCAATGAAATGCAACCTTTATGTCATGC 188
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAAGACGGGATTGTGAGG 216

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-470-124-47

seq_documentation_block:
Sequence 47, Application US/09470124
GENERAL INFORMATION:
APPLICANT: Mason
APPLICANT: Arntzen
TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
TITLE OF INVENTION: Transgenic Plants
FILE REFERENCE: 4868/84454
CURRENT APPLICATION NUMBER: US/09/470,124
CURRENT FILING DATE: 1999-12-24
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 47
LENGTH: 777
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(777)
OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
OTHER INFORMATION: cholera toxin gene mutagenized to optimize
OTHER INFORMATION: expression in plants.
US-09-470-124-47

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-470-124-47 ..

Align seg 1/1 to: US-09-470-124-47 from: 1 to: 777

29 GluTyrPheaspArgGlyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
139 GAGTACTTTGACGAGGCTACTCAATGAAATGCAACCTTTATGTCATGC 188
45 aArgGlyThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAACTCAAGACGGGATTGTGAGG 216

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-470-124-49


```

seq_documentation_block:
; Sequence 49, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-49

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-470-124-49 ..
Align seg 1/1 to: US-09-470-124-49 from: 1 to: 777

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTGTGACGGGTACTCAGATGACATGCAACCTTTATGACCATGC 188
|||||
45 AARGGLYThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAGACTCAACCTGATTTGTGAGG 216

seq_name: /cgn2_6/ptodata/2/pna/US094_COMB.seq:us-09-470-124-51

seq_documentation_block:
; Sequence 51, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed in
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:V. cholerae
; OTHER INFORMATION: cholera toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-51

```

```

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-470-124-51 ..
Align seg 1/1 to: US-09-470-124-51 from: 1 to: 777

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
139 GAGTACTTGTGACGGGTACTCAGATGACATGCAACCTTTATGACCATGC 188
|||||
45 AARGGLYThrGlnThrGlyPheValArg 54
|||||
189 AAGGGAGACTCAACCTGATTTGTGAGG 216

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US97-11719-12

seq_documentation_block:
; Sequence 12, Application PC/TUS9711719
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IMMUNOCONTRACEPTION COMPOSITIONS
; TITLE OF INVENTION: CONTAINING SPERM ANTIGEN, AND METHODS OF USE
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/11719
; FILING DATE: 03-JUL-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/674,895
; FILING DATE: 03-JUL-1996
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1956 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; PCT-US97-11719-12

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US97-11719-12 ..
Align seg 1/1 to: PCT-US97-11719-12 from: 1 to: 1956

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
1053 GAGTACTTGTGACGGGTACTCAGATGACATGCAACCTTTATGATCATGC 1102
|||||
45 AARGGLYThrGlnThrGlyPheValArg 54
|||||
1103 AAGAGAGACTCAGACGGGATTTGTTAGG 1130

seq_name: /cgn2_6/ptodata/2/pna/US086_COMB.seq:us-08-674-895-12

seq_documentation_block:
; Sequence 12, Application US/08674895
; GENERAL INFORMATION:

```


APPLICANT: Blell, Jeffrey D.
APPLICANT: Bookbinder, Louis H.
APPLICANT: Helm, Mich B.
TITLE OF INVENTION: IMMUNOCONTRACEPTION COMPOSITIONS
TITLE OF INVENTION: CONTAINING SPERM ANTIGEN, AND METHODS OF USE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 North Torrey Pines Road
CITY: La Jolla
STATE: California
COUNTRY: United States
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/674,895
FILING DATE: 03-JUN-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 529.0
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1956 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-674-895-12

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-674-895-12 ..

Align seg 1/1 to: US-08-674-895-12 from: 1 to: 1956

seq_name: /cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq:PCT-US01-08582-1

seq_documentation_block:
Sequence 1, Application PC/TUS0108582
GENERAL INFORMATION:
APPLICANT: UAB Research Foundation
TITLE OF INVENTION: Chimeric Nontoxic Mutants of
Enterotoxins as Mucosal Adjuvants for Cell-Mediated or
Humoral Immunity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hendricks and Assoc
STREET: P.O. Box 2509
CITY: Fairfax

29 GluTyPheAspArgGlyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
1053 GAGTACTTGTGACCGAGGTACTCAATGAATGATCAACCTTATGATCATGC 1102
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
1103 AAGAGGAACTCAGACGCGGATTGTGTAGG 1130

STATE: VA
COUNTRY: US
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US01/08582
FILING DATE: 16-Mar-2001
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: MCG-01
TELEPHONE: 703/425-8405
TELEFAX: 703/425-8406
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2022 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
PCT-US01-08582-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US01-08582-1 ..

Align seg 1/1 to: PCT-US01-08582-1 from: 1 to: 2022

seq_name: /cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq:PCT-US01-08582-2

seq_documentation_block:
Sequence 2, Application PC/TUS0108582
GENERAL INFORMATION:
APPLICANT: UAB Research Foundation
TITLE OF INVENTION: Chimeric Nontoxic Mutants of
Enterotoxins as Mucosal Adjuvants for Cell-Mediated or
Humoral Immunity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hendricks and Assoc
STREET: P.O. Box 2509
CITY: Fairfax
STATE: VA
COUNTRY: US
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

29 GluTyPheAspArgGlyThrGlnMetAsnIleasnLeuTyrAspHisAl 45
|||||
656 GAGTACTTGTGACCGAGGTACTCAATGAATGATCAACCTTATGATCATGC 705
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
706 AAGAGGAACTCAGACGCGGATTGTGTAGG 733

APPLICATION NUMBER: PCT/US01/08582
FILING DATE: 16-Mar-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: MCG-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703/425-8405
TELEFAX: 703/425-8406
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2022 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
PCT-US01-08582-2

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US01-08582-2 ..

Align seq 1/1 to: PCT-US01-08582-2 from: 1 to: 2022

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
656 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTATGATCATGC 705
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
706 AAGAGGAACTCAGACGGGATTGTAGG 733

seq_name: /cgn2_6/ptodata/2/pna/US097B.COMB.seq:US-09-724-315-3

seq_documentation_block:

Sequence 3, Application US/09724315
GENERAL INFORMATION:
APPLICANT: HAYNES, Joel R.
TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
FILE REFERENCE: APF41
CURRENT APPLICATION NUMBER: US/09/724,315
CURRENT FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 5488
TYPE: DNA
ORGANISM: plasmid pPV2006
US-09-724-315-3

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-724-315-3 ..

Align seq 1/1 to: US-09-724-315-3 from: 1 to: 5488

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||

4059 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTATGATCATGC 4108
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
4109 AAGAGGAACTCAGACGGGATTGTAGG 4136

seq_name: /cgn2_6/ptodata/2/pna/US097B.COMB.seq:US-09-724-315-1

seq_documentation_block:

Sequence 1, Application US/09724315
GENERAL INFORMATION:
APPLICANT: HAYNES, Joel R.
TITLE OF INVENTION: NUCLEIC ACID ADJUVANTS
FILE REFERENCE: APF41
CURRENT APPLICATION NUMBER: US/09/724,315
CURRENT FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 5500
TYPE: DNA
ORGANISM: pPV2002 plasmid
US-09-724-315-1

alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-724-315-1 ..

Align seq 1/1 to: US-09-724-315-1 from: 1 to: 5500

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
4059 GAGTACTTTGACCGAGGTACTCAATGATATCAACCTTATGATCATGC 4108
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
4109 AAGAGGAACTCAGACGGGATTGTAGG 4136

seq_name: /cgn2_6/ptodata/2/pna/US086.COMB.seq:US-08-670-974-6

seq_documentation_block:

Sequence 6, Application US/08670974
GENERAL INFORMATION:
APPLICANT: Mekalanos, John J.
TITLE OF INVENTION: NOVEL CHOLERA VECTORS, VACCINES,
TITLE OF INVENTION: AND METHODS FOR ANTIGEN DELIVERY IN GRAM-NEGATIVE
TITLE OF INVENTION: BACTERIA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,974
FILING DATE: 26-JUN-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bieker-Brady, Kristina
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 00742/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 6943 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
US-08-670-974-6

Alignment_scores:
Quality: 26.00 Length: 26
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-08-670-974-6 ..

Align seg 1/1 to: US-08-670-974-6 from: 1 to: 6943

29 GIUTYrPhaSPaRgGlyThrGlnMetAsnIleAsnLeuTyraSPHsAl 45
|||||
5829 GAGTACTTTGACCGAGGTACTCAATGATATACCTTTATATCATGTC 5878
|||||
45 AARGGLYThrGlnThrclyPheValArg 54
|||||
5879 AAGAGGAACTCAGACGAGATTGTTAGG 5906

seq_name: /cgn2_6/ptodata/2/pna/PCrUS_COMB.seq:PCT-US01-16904-1

seq_documentation_block:
; Sequence 1, Application PC/PTUS0116904
; GENERAL INFORMATION:
; APPLICANT: University of Maryland Biotechnology Institute
; APPLICANT: Powell, Robert J.
; APPLICANT: Hone, David M.
; TITLE OF INVENTION: Methods for Introducing and Expressing Genes in Animal Cells, and
; TITLE OF INVENTION: Bacterial Bilebs for Use in Same
; FILE REFERENCE: 4115-121 PCT
; CURRENT APPLICATION NUMBER: PCT/US01/16904
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/206,994
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 7586
; TYPE: DNA
; ORGANISM: Cholera Toxin A1 Subunit Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6184)..(6184)
; OTHER INFORMATION: n can be any one of a, c, g, and t.
; NAME/KEY: misc_feature
; LOCATION: (6224)..(6224)
; OTHER INFORMATION: n can be any one of a, c, g, and t.
; PCT-US01-16904-1

alignment_scores:
Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US01-16904-1 ..

Align seg 1/1 to: PCT-US01-16904-1 from: 1 to: 7586

29 GIUTYrPhaSPaRgGlyThrGlnMetAsnIleAsnLeuTyraSPHsAl 45
|||||
87 GAGTACTTTGACCGAGGTACTCAATGATATACCTTTATATCATGTC 136
|||||
45 AARGGLYThrGlnThrclyPheValArg 54
|||||
137 AAGAGGAACTCAGACGAGATTGTTAGG 164

seq_name: /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-040-990-1

seq_documentation_block:
; Sequence 1, Application US/09040990
; GENERAL INFORMATION:
; APPLICANT: Clements, J.
; APPLICANT: Friede, M.
; TITLE OF INVENTION: USE OF MUTANT ENTEROTOXIN WITH EXCESS B-SUBUNIT AS AN
; TITLE OF INVENTION: ADJUVANT
; FILE REFERENCE: 5113-052
; CURRENT APPLICATION NUMBER: US/09/040,990
; CURRENT FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(45)
US-09-040-990-1

alignment_scores:
Quality: 15.00 Length: 15
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-040-990-1 ..

Align seg 1/1 to: US-09-040-990-1 from: 1 to: 45

186 GLYcysGlyAsnSerSerArgThrIleThrclyAspThrcysAsn 200
|||||
1 GGTGTGGAATTCATCAAGACAAATTACAGTGATCTGTGTAAT 25

seq_name: /cgn2_6/ptodata/2/pna/US093_COMB.seq:US-09-365-530-1

seq_documentation_block:
; Sequence 1, Application US/09365530
; GENERAL INFORMATION:
; APPLICANT: Clements, John D.
; APPLICANT: Dickinson, Bonny L.
; TITLE OF INVENTION: MUTANT ENTEROTOXIN EFFECTIVE AS A
; TITLE OF INVENTION: NON-TOXIC ORAL ADJUVANT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/365,530
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/296,848
FILING DATE: 26-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5113-046
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-09-365-530-1

```

```

alignment_scores:
Quality: 15.00      Length: 15
Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x US-09-365-530-1 ..

```

```

Align seg 1/1 to: US-09-365-530-1 from: 1 to: 45

```

```

186 GLYCYSGLYASNSERFARGTHTLRGLVSPTRHCYSASN 200
|||||
1 GGTGTGGAATTCATCAAGACAATTCACAGTGCTACTTGTAA 45

```

```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-8

```

```

seq_documentation_block:
; Sequence 8, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LT-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-8

```

```

alignment_scores:
Quality: 13.00      Length: 13
Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x PCT-US99-30747-8 ..

```

```

Align seg 1/1 to: PCT-US99-30747-8 from: 1 to: 40

```

```

19 SERGLYGLYLEUWETPROARGGLYHISASNGLTYPHE 31
|||||
2 TCTGAGCTCTCATGCCAGGACACAAATGACTACTT 40

```

```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-9

```

```

seq_documentation_block:
; Sequence 9, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LT-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-9

```

```

alignment_scores:
Quality: 13.00      Length: 13
Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x PCT-US99-30747-9 ..

```

```

Align seg 1/1 to: PCT-US99-30747-9 from: 1 to: 40

```

```

32 ASPARGLYTHRGLMETASNLEASNLEUTYRASPHIS 44
|||||
1 GATAGGGGAGACCCAAATGACATCACTCTATGACCAT 39

```

```

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-11

```

```

seq_documentation_block:
; Sequence 11, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LT-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-11

```

```

alignment_scores:
Quality: 13.00      Length: 13
Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```


alignment_block:
US-09-528-682-1 x PCT-US99-30747-11 ..

Align seg 1/1 to: PCT-US99-30747-11 from: 1 to: 40

59 TyrValSerThrSerLeuSerLeuArgSerAlaHisLeu 71
|||||
2 TATGTGTCCACCTCCTCTAGCTTGAGTGTCTGCCTCCTG 40

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-12

seq_documentation_block:

; Sequence 12, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 12
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-12

alignment_scores:

Quality:	13.00	Length:	13
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-12 ..

Align seg 1/1 to: PCT-US99-30747-12 from: 1 to: 40

72 AlaGlyGlnSerIleuSerGlyTyrSerThrTyrTyr 84
|||||
1 GCAGGACAATCCATCCTCTCAGGATACCTCCACTACTAC 39

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-14

seq_documentation_block:

; Sequence 14, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 14
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-14

alignment_scores:

Quality:	13.00	Length:	13
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-14 ..

Align seg 1/1 to: PCT-US99-30747-14 from: 1 to: 40

99 AspValLeuGlyValTyrSerProHisProTyrGln 111
|||||
2 GATGTGTGGAGAGTATAGCCCTCACCACCATATGAGCAA 40

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-15

seq_documentation_block:

; Sequence 15, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 15
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-15

alignment_scores:

Quality:	13.00	Length:	13
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-15 ..

Align seg 1/1 to: PCT-US99-30747-15 from: 1 to: 40

112 GluValSerAlaLeuGlyIleProTyrSerGln 124
|||||
1 GAGGTGTCTGCTTGGGTGCANTCCTACTCCCAATTC 39

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-17

seq_documentation_block:

; Sequence 17, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 17
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-17


```
alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-17  ..

Align seg 1/1 to: PCT-US99-30747-17 from: 1 to: 40

139 LeuHISarGAsnArGluTYrArGAspArGTYrTYrArG 151
|||||
2 CTCATAGAGATAGGAGATATAGGACAGGATCTATAGG 40

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-18

seq_documentation_block:
; Sequence 18, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 18
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lf-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-18

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-18  ..

Align seg 1/1 to: PCT-US99-30747-18 from: 1 to: 40

152 AsnLeuAsnIleAlaProAlaGluAspGlyTYrArGleu 164
|||||
1 AACCTCAACATAGCTCCAGCAGAGAGATGTTATAGTTG 39

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-21

seq_documentation_block:
; Sequence 21, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 21
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
```

```
; OTHER INFORMATION: used to assemble Lf-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-21

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-21  ..

Align seg 1/1 to: PCT-US99-30747-21 from: 1 to: 40

192 ArgThrIleThrcGlyAspThrcGysAsnGluGluThrcGln 204
|||||
1 AGGACATCAAGAGTGACCTTGCAATGAGAGACCA 39

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-23

seq_documentation_block:
; Sequence 23, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 23
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lf-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-23

alignment_scores:
  Quality: 13.00      Length: 13
  Ratio: 1.000        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-528-682-1 x PCT-US99-30747-23  ..

Align seg 1/1 to: PCT-US99-30747-23 from: 1 to: 40

219 LysArgGlnIlePheSerAspTYrGlnSerGluValAsp 231
|||||
2 AAGAGCAAAATCTTCCAGACTACCAATCAGAGGTGGAC 40

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-28

seq_documentation_block:
; Sequence 28, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 28
; LENGTH: 40
; TYPE: DNA
```



```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LTR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-28

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-28/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-28 from: 1 to: 40

199 CysAnGluGluThrGlnAsnLeuSerThrIleTyrLeu 211
|||||
39 TGCAATGAGGAGACCCAAACCTTAGCACATCTACCTT 1

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-30

seq_documentation_block:
; Sequence 30, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LTR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-30

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-30/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-30 from: 1 to: 40

172 GlnAlaThrParGluGluProTyrPheHisHisAlaPro 184
|||||
40 CAAGCCTGGAGGAGGAGCCCTGATCCACCATGCACCA 2

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-31

seq_documentation_block:
; Sequence 31, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 31
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LTR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-31

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-31/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-31 from: 1 to: 40

159 GluAspGlyTyrArgLeuAlaGlyPheProPronAspHis 171
|||||
39 GAGGATGTTATAGTTGCGAGTTTCACACAGACGAC 1

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-33

seq_documentation_block:
; Sequence 33, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LTR-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-33

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-33/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-33 from: 1 to: 40

133 PheGlyValIleAspArgLeuHisArgAsnArgGlu 144
|||||
40 TTTCGTGATGATTGATGAGAGCGCTCATAGAGATAGCGAG 2

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-34

seq_documentation_block:
; Sequence 34, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
```



```
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lr-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-34

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-34/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-34 from: 1 to: 40

119 ILEPRTYISERGLIETRYGIYTPTRYARVALASN 131
|||||
39 ATCCCATCTCCCAATCTATGATGTGTATAGGTGAC 1

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-36

seq_documentation_block:
; Sequence 36, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lr-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-36

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-36/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-36 from: 1 to: 40

92 PROASNMEPTHASNYVALASNPVALLEUGLYVALTYR 104
|||||
40 CCAACATGTTCAATGATGATGTGTGGAGTGTAT 2

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-37

seq_documentation_block:
; Sequence 37, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
```

```
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lr-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-37

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-37/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-37 from: 1 to: 40

79 GLYTYRSETRHTYRTRYILETYRVALILEALATHRALA 91
|||||
39 GGATFACTCCACCTACTACATCTATGATGTGCTACACCA 1

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-39

seq_documentation_block:
; Sequence 39, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble Lr-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-39

alignment_scores:
    Quality: 13.00      Length: 13
    Ratio: 1.000        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US99-30747-39/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-39 from: 1 to: 40

52 PHEVALARGTYRSPASPGIYTYRVALSERTHSERLEU 64
|||||
40 TTTGTGAGTATGATGATGATGTGTCCACCTCCCTT 2

seq_name: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:PCT-US99-30747-40

seq_documentation_block:
; Sequence 40, Application PC/TUS9930747
```



```

; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
; OTHER INFORMATION: used to assemble LT-A gene for expression in
; OTHER INFORMATION: plants.
PCT-US99-30747-40

```

```

alignment_scores:
      Quality: 13.00      Length: 13
      Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

```

alignment_block:

US-09-528-682-1 x PCT-US99-30747-40/rev ..

Align seg 1/1 to reverse of: PCT-US99-30747-40 from: 1 to: 40

```

39 IleasnleuTYrAspHisAlaArgGlyThrGlnThrGly 51
|||||
39 ATCACCTCTATGACCATGCTAGGGCAACCAACTGTC 1

```


OM of: US-09-528-682-1 to: Pending_Patents_NA_New:* out_format : pfs
Date: Jun 18, 2002 11:55 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-Q/cg2n_1/USPRO.spool/US09528682/runat_18062002.082607.8882/app-query.fasta_1.299
-Db=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=olip2n.rnpn
-GAPOP=4.500 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPECL=0.000
-LOOPEXT=0.000 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=60.000
-XGAPEXT=60.000 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000
-YGAPEXT=60.000 -DELOP=6.000 -DELEXT=7.000 -START=1
-MATRIX=oligo -TRANS=numan40.cdi -LIST=1000 -DOCALIGN=200
-THR_SCORE=quality -THR_MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09528682.@CGN1_1.476 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPRX -WAIT -THREADS=1

Search information block:

Query: US-09-528-682-1
Query length: 240
Database: Pending_Patents_NA_New:*
Database sequences: 1014543
Search time (sec): 224.630000

MARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
MARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-809-033A-1 +	26.00	477.73	1.7e-18	20	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-809-033A-2 +	26.00	477.73	1.7e-18	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-105-299-13399 -	8.00	137.84	14.52	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-24256 +	8.00	134.54	22.09	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-67507 -	8.00	134.10	23.46	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-8851 +	8.00	132.84	27.59	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-203067 -	8.00	132.84	27.59	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-8852 +	8.00	132.00	30.71	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-125010 +	8.00	131.38	33.26	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-333968 +	8.00	125.27	72.84	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-258165 -	8.00	123.70	89.09	20	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-112146 +	7.00	139.92	11.12	22	
/cg2n_6/pdataa/1/pna/US10_NEW_COMB.seq:US-10-027-633-3331D-20655 -	7.00	129.23	43.82	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-20655 -	7.00	128.54	47.84	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-975-234-14267 +	7.00	122.94	98.12	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-673-476-266 +	7.00	122.54	103.28	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8128 +	7.00	122.23	105.42	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8128 +	7.00	122.23	105.42	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-13594 +	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25	141.62	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.11	141.15	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11333 +	7.00	121.61	113.35	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235 -	7.00	121.27	121.65	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360 -	7.00	121.08	125.90	22	
/cg2n_6/pdataa/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303 +	7.00	120.25			

[illegible][illegible]

[illegible]

/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-142-766-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-139-980-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-274-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-471-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-473-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-806-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-807-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-808-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-809-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-810-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-861-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-863-409 + 7.00 102.07 1.4e+03 3
/cgn2_6/ptodata/1/pna/US10_NEW.COMB.seq:US-10-140-921-409 + 7.00 102.07 1.4e+03 3

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW.COMB.seq:US-09-809-033A-1

seq_documentation_block:

; Sequence 1, Application US/09809033A

; GENERAL INFORMATION:

; APPLICANT: McGhee, Jerry

; Kiyono, Hiroshi

; Takeda, Yoshifumi

; Ohmura, Mari

; Yamamoto, Shingo

; TITLE OF INVENTION: Chimeric Nontoxic Mutants of

; Enterotoxins as Mucosal Adjuvants for Cell-Mediated or

; Humoral Immunity

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hendricks and Assoc

; STREET: P.O. Box 2509

; CITY: Fairfax

; STATE: VA

; COUNTRY: US

; ZIP: 22031

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/809,033A

; FILING DATE: 16-Mar-2001

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Hendricks, Glena

; REGISTRATION NUMBER: 32,535

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703/425-8405

; TELEFAX: 703/425-8406

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2022 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-809-033A-1

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-809-033A-1 ..

Align seg 1/1 to: US-09-809-033A-1 from: 1 to: 2022

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
656 GAGTACTTGGACCGAGGTACTCAATGATATATCACTTATGATCATGCG 705
45 aargGlyThrGlnThrGlyPheValArg 54
|||||
706 AAGAGAACTGACGAGGTACTTGTATGAG 733

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW.COMB.seq:US-09-809-033A-2

seq_documentation_block:

; Sequence 2, Application US/09809033A

; GENERAL INFORMATION:

; APPLICANT: McGhee, Jerry

; Kiyono, Hiroshi

; Takeda, Yoshifumi

; Ohmura, Mari

; Yamamoto, Shingo

; TITLE OF INVENTION: Chimeric Nontoxic Mutants of

; Enterotoxins as Mucosal Adjuvants for Cell-Mediated or

; Humoral Immunity

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hendricks and Assoc

; STREET: P.O. Box 2509

; CITY: Fairfax

; STATE: VA

; COUNTRY: US

; ZIP: 22031

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/809,033A

; FILING DATE: 16-Mar-2001

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Hendricks, Glena

; REGISTRATION NUMBER: 32,535

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703/425-8405

; TELEFAX: 703/425-8406

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2022 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-809-033A-2

alignment_scores:

Quality: 26.00 Length: 26

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-809-033A-2 ..

Align seg 1/1 to: US-09-809-033A-2 from: 1 to: 2022

29 GluTyrPheAspArgGlyThrGlnMetAsnIleAsnLeuTyrAspHisAl 45
|||||
656 GAGTACTTGGACCGAGGTACTCAATGATATATCACTTATGATCATGCG 705

45 AAGGAGTACGAGGCGGATTGTAGG 54
 ||||||||||||||||||
 706 AAGGAGTACGAGGCGGATTGTAGG 733

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-105-299-12399

```
seq_documentation_block:
; Sequence 12399, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12399
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12399
```

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-105-299-12399/rev ..

Align seg 1/1 to reverse of: US-10-105-299-12399 from: 1 to: 406

75 SerIleuSerGlyrSerThr 82
 ||||||||||||||||||
 100 TCATCCTCTGAGTGTACACACACA 77

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-242566

```
seq_documentation_block:
; Sequence 242566, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 242566
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-242566
```

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-67507/rev ..

Align seg 1/1 to reverse of: US-10-027-632-67507 from: 1 to: 651

61 SerThrSerIleuSerArgSer 68
 ||||||||||||||||||
 390 TCACACTCTCTGAGCCTCTGATCC 367

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-203067

```
seq_documentation_block:
; Sequence 203067, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203067
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-203067
```

Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-242566 ..

Align seg 1/1 to: US-10-027-632-242566 from: 1 to: 631

119 IleProTySerGlnIleTyGly 126
 ||||||||||||||||||
 383 ATCTCTACTCTCAGATCTATGGA 406

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-67507

```
seq_documentation_block:
; Sequence 67507, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67507
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-67507
```

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-67507/rev ..

Align seg 1/1 to reverse of: US-10-027-632-67507 from: 1 to: 651

61 SerThrSerIleuSerArgSer 68
 ||||||||||||||||||
 390 TCACACTCTCTGAGCCTCTGATCC 367

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-203067

```
seq_documentation_block:
; Sequence 203067, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203067
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-203067
```

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-67507/rev ..

Align seg 1/1 to reverse of: US-10-027-632-67507 from: 1 to: 651

61 SerThrSerIleuSerArgSer 68
 ||||||||||||||||||
 390 TCACACTCTCTGAGCCTCTGATCC 367

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-203067

```
seq_documentation_block:
; Sequence 203067, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203067
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-203067
```



```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1998-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203067
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-203067
```

```
alignment_scores:
    Quality: 8.00      Length: 8
    Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-203067/rev ..

Align seg 1/1 to reverse of: US-10-027-632-203067 from: 1 to: 672

```
113 ValSerAlaLeuGlyGlyIlePro 120
|||||
645 GTATCTGCTTAGGGGCGCATCCCA 622
```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-8851

```
seq_documentation_block:
; Sequence 8851, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1998-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8851
; LENGTH: 797
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; NAME/KEY: misc_feature
```

```
; LOCATION: (1)...('797)
; OTHER INFORMATION: n = A,T,C or G
; US-10-027-632-8851
```

```
alignment_scores:
    Quality: 8.00      Length: 8
    Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-8851/rev ..

Align seg 1/1 to reverse of: US-10-027-632-8851 from: 1 to: 797

```
68 SerAlaHisLeuAlaGlyGlnSer 75
|||||
277 AGTCCCATCTGCTGCGCAGAGC 254
```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-8852

```
seq_documentation_block:
; Sequence 8852, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8852
; LENGTH: 797
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...('797)
; OTHER INFORMATION: n = A,T,C or G
; US-10-027-632-8852
```

```
alignment_scores:
    Quality: 8.00      Length: 8
    Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

alignment_block:

US-09-528-682-1 x US-10-027-632-8852/rev ..

Align seg 1/1 to reverse of: US-10-027-632-8852 from: 1 to: 797

```
68 SerAlaHisLeuAlaGlyGlnSer 75
|||||
277 AGTCCCATCTGCTGCGCAGAGC 254
```

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-125010


```
seq_documentation_block:
; Sequence 125010 Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125010
; LENGTH: 892
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-125010

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-125010 ..

Align seg 1/1 to: US-10-027-632-125010 from: 1 to: 892
65 SerLeuArgSerAlaHisLeuAla 72
|||||
131 TCCTTGAGAGTCCCATCTGCC 154

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-323968

seq_documentation_block:
; Sequence 323968 Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 323968
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-323968

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-323968 ..

Align seg 1/1 to: US-10-027-632-323968 from: 1 to: 970
65 SerLeuArgSerAlaHisLeuAla 72
|||||
221 TCCTTGAGAGTCCCATCTGCC 244

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-258164

seq_documentation_block:
; Sequence 258164 Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258164
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258164

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-258164/rev ..

Align seg 1/1 to reverse of: US-10-027-632-258164 from: 1 to: 2211
150 TyTArGAsnLeuAsnIleAlaPro 157
|||||
819 TACAGGAATCTGAATATTGCTCCG 796

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-258165

seq_documentation_block:
; Sequence 258165 Application US/10027632
```

```
; SEQ ID NO 323968
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-323968

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-323968 ..

Align seg 1/1 to: US-10-027-632-323968 from: 1 to: 970
65 SerLeuArgSerAlaHisLeuAla 72
|||||
221 TCCTTGAGAGTCCCATCTGCC 244

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-258164

seq_documentation_block:
; Sequence 258164 Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258164
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258164

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-258164/rev ..

Align seg 1/1 to reverse of: US-10-027-632-258164 from: 1 to: 2211
150 TyTArGAsnLeuAsnIleAlaPro 157
|||||
819 TACAGGAATCTGAATATTGCTCCG 796

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-258165

seq_documentation_block:
; Sequence 258165 Application US/10027632
```



```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258165
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-258165

alignment_scores:
    Quality: 8.00      Length: 8
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-258165/rev ..

Align seg 1/1 to reverse of: US-10-027-632-258165 from: 1 to: 2211

150 TTTATGAsnleuAsnllleAPro 157
|||||
819 TACAGGATCTGAAATGCTCG 796 .

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-112146

seq_documentation_block:
; Sequence 112146, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112146
; LENGTH: 2732
```

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-112146

alignment_scores:
    Quality: 8.00      Length: 8
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-112146 ..

Align seg 1/1 to: US-10-027-632-112146 from: 1 to: 2732

162 TTTATGleuAlaglypHePro 169
|||||
193 TACAGGCTTGACAGGTCCACCT 216

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-138-381-26

seq_documentation_block:
; Sequence 26, Application US/10138381
; GENERAL INFORMATION:
; APPLICANT: NAKAYAMA, Tomoko
; APPLICANT: TADA, Jun
; APPLICANT: FUKUSHIMA, Shigeru
; APPLICANT: OHASHI, Tetsuo
; TITLE OF INVENTION: Oligonucleotides for Detecting Bacteria and
; FILE REFERENCE: 1422-0430P
; CURRENT APPLICATION NUMBER: US/10/138,381
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US/09/614,681
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 08/968,046
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 08/328,710
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: 6-48174 JAPAN
; PRIOR FILING DATE: 1994-03-18
; PRIOR APPLICATION NUMBER: 6-30277 JAPAN
; PRIOR FILING DATE: 1994-02-28
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Vibrio cholerae
US-10-138-381-26

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-138-381-26/rev ..

Align seg 1/1 to reverse of: US-10-138-381-26 from: 1 to: 22

122 SerGlnleTyrGlyTTPtyr 128
|||||
21 TCCCAATATATGATGTTAT 1

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-53058

seq_documentation_block:
; Sequence 53058, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```


;; TITLE OF INVENTION: Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 53058
;; LENGTH: 93
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-53058

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-53058/rev ..

Align seg 1/1 to reverse of: US-10-027-632-53058 from: 1 to: 93

66 leuArgSerIaHsleuA 72
|||||
27 TTAGATCCGACATTTAGCG 7

seq_name: /cgn2_6/ptodata/1/pna/US09_NEM_COMB.seq:US-09-539-331D-20665

seq_documentation_block:

;; Sequence 20665, Application US/09539331D
;; GENERAL INFORMATION:
;; APPLICANT: Seilhamer, Jeffrey J.
;; APPLICANT: Delegeane, Angelo M.
;; APPLICANT: Stuart, Susan G.
;; APPLICANT: Stuve, Laura L.
;; APPLICANT: Mulahy, Sara J.
;; APPLICANT: Naughton, Rebecca E.
;; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
;; FILE REFERENCE: PD-1022 CIP
;; CURRENT APPLICATION NUMBER: US/09/539,331D
;; CURRENT FILING DATE: 2000-03-30
;; PRIOR APPLICATION removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 40961
;; SOFTWARE: PERL Program
;; SEQ ID NO: 20665
;; LENGTH: 102
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; OTHER INFORMATION: Incyte ID No: hu00919904
US-09-539-331D-20665

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-539-331D-20665/rev ..

Align seg 1/1 to reverse of: US-09-539-331D-20665 from: 1 to: 102

63 SerIeuSerIeuArgSerIa 69
|||||
29 AGCCTGAGTCTGAGAGTGC 9

seq_name: /cgn2_6/ptodata/1/pna/US09_NEM_COMB.seq:US-09-975-254-14267

seq_documentation_block:

;; Sequence 14267, Application US/09975254
;; GENERAL INFORMATION:
;; APPLICANT: Byrum, Joseph R.
;; APPLICANT: Heck, Gregory R.
;; APPLICANT: La Rosa, Thomas J.
;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;; FILE REFERENCE: 38-21(15309)B
;; CURRENT APPLICATION NUMBER: US/09/975,254
;; CURRENT FILING DATE: 2001-10-12
;; PRIOR APPLICATION NUMBER: US/09/263,191
;; PRIOR FILING DATE: 1999-03-05
;; NUMBER OF SEQ ID NOS: 31255
;; SEQ ID NO: 14267
;; LENGTH: 158
;; TYPE: DNA
;; ORGANISM: Glycine max
;; OTHER INFORMATION: Clone ID: 700908892H1
US-09-975-254-14267

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-975-254-14267 ..

Align seg 1/1 to: US-09-975-254-14267 from: 1 to: 158

63 SerIeuSerIeuArgSerIa 69
|||||
7 TCCCTCTCTGCGCTGCGCA 27

seq_name: /cgn2_6/ptodata/1/pna/US09_NEM_COMB.seq:US-09-673-476-266

seq_documentation_block:

;; Sequence 266, Application US/09673476
;; GENERAL INFORMATION:
;; APPLICANT: COLE, STEWART
;; APPLICANT: BUCHRIESER-BROSCH, ROLAND
;; APPLICANT: GORDON, STEPHEN
;; APPLICANT: BILLAULT, ALAIN
;; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
;; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
;; TITLE OF INVENTION: DNA LIBRARY, APPLICATION TO THE DETECTION OF
;; FILE REFERENCE: 05394.0011-00000
;; CURRENT APPLICATION NUMBER: US/09/673,476
;; CURRENT FILING DATE: 2002-03-29
;; PRIOR APPLICATION NUMBER: PCT/IB99/00740
;; PRIOR FILING DATE: 1999-04-16
;; PRIOR APPLICATION NUMBER: 09/060,756
;; PRIOR FILING DATE: 1998-04-16
;; NUMBER OF SEQ ID NOS: 743
;; SOFTWARE: PatentIn Ver. 2.2
;; SEQ ID NO: 266
;; LENGTH: 217
;; TYPE: DNA


```
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (139)..(140)
; OTHER INFORMATION: a, t, c or g
US-09-673-476-266
```

alignment_scores:

```
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-09-673-476-266 ..
```

```
Align seg 1/1 to: US-09-673-476-266 from: 1 to: 217
```

```
63 SerleusSerleuArgSerAla 69
|||||
10 TCCTATCGCTCGCTCTGCA 30
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8128
```

seq_documentation_block:

```
; Sequence 8128, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara L.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 8128
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00121720
; FEATURE:
```

```
; NAME/KEY: unsure
; LOCATION: 60, 178
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-800C-8128
```

alignment_scores:

```
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-09-539-800C-8128 ..
```

```
Align seg 1/1 to: US-09-539-800C-8128 from: 1 to: 229
```

```
63 SerleusSerleuArgSerAla 69
|||||
92 TCACCTTCGCTCGCTCTGCT 112
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-975-254-23834
```

seq_documentation_block:

```
; Sequence 23834, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and other Molecules Associated With
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 23834
; LENGTH: 234
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700962842H1
US-09-975-254-23834
```

alignment_scores:

```
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

```
US-09-528-682-1 x US-09-975-254-23834/rev ..
```

```
Align seg 1/1 to reverse of: US-09-975-254-23834 from: 1 to: 234
```

```
63 SerleusSerleuArgSerAla 69
|||||
173 TCCTAAGCTACGCTCTGCT 153
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-13594
```

seq_documentation_block:

```
; Sequence 13594, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara L.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
```



```
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 13594
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00527775
; NAME/KEY: unsure
; LOCATION: 206
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-331D-13594

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-539-331D-13594 ..
Align seg 1/1 to: US-09-539-331D-13594 from: 1 to: 239

153 LeuAsn11eAlaProAlaGlu 159
|||||
119 CTCACATGTCACGACGACAG 139

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-11353
seq_documentation_block:
; Sequence 11353, Application US/09540210B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526
; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
```

```
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 11353
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00067358
; NAME/KEY: unsure
; LOCATION: 7, 93, 186, 204
; OTHER INFORMATION: a, t, c, g, or other
```


US-09-540-210B-11353

alignment_scores:

Quality: 7.00 Length: 7
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-540-210B-11353

Align seg 1/1 to: US-09-540-210B-11353 from: 1 to: 242

217 LysValLysArgGlnIlePhe 223
 |||||
 143 AAGGTAAAGACAGATATT 163

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-9255

seq_documentation_block:

; Sequence 9255, Application US/09539331D
 ; GENERAL INFORMATION:
 ; APPLICANT: Seilhamer, Jeffrey J.
 ; APPLICANT: Deleageane, Angelo M.
 ; APPLICANT: Stuart, Susan G.
 ; APPLICANT: Stuve, Laura L.
 ; APPLICANT: Mullahy, Sara J.
 ; APPLICANT: Naughton, Rebecca E.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
 ; FILE REFERENCE: PD-1022 CIP
 ; CURRENT APPLICATION NUMBER: US/09/539,331D
 ; CURRENT FILING DATE: 2000-03-30
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 40961
 ; SOFTWARE: PERL Program
 ; SEQ ID NO: 9255
 ; LENGTH: 253
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No: hu00523285
 ; US-09-539-331D-9255

alignment_scores:

Quality: 7.00 Length: 7
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-539-331D-9255/rev

Align seg 1/1 to reverse of: US-09-539-331D-9255 from: 1 to: 253

66 LeuArgSerAlaHisLeuAla 72
 |||||
 119 TTACGGCTCTCACACCTGGCC 99

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-8235

seq_documentation_block:

; Sequence 8235, Application US/09539800C
 ; GENERAL INFORMATION:
 ; APPLICANT: Seilhamer, Jeffrey J.
 ; APPLICANT: Deleageane, Angelo M.
 ; APPLICANT: Stuart, Susan G.
 ; APPLICANT: Stuve, Laura L.
 ; APPLICANT: Mullahy, Sara J.
 ; APPLICANT: Naughton, Rebecca E.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
 ; FILE REFERENCE: PD-1023 CIP
 ; CURRENT APPLICATION NUMBER: US/09/539,800C

; CURRENT FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: 08/521,383
 ; PRIOR FILING DATE: August 16, 1995
 ; PRIOR APPLICATION NUMBER: 08/271,217
 ; PRIOR FILING DATE: June 27, 1994
 ; PRIOR APPLICATION NUMBER: 08/334,881
 ; PRIOR FILING DATE: November 4, 1994
 ; PRIOR APPLICATION NUMBER: 08/943,978
 ; PRIOR FILING DATE: October 3, 1997
 ; PRIOR APPLICATION NUMBER: 60/028,732
 ; PRIOR FILING DATE: October 4, 1996
 ; PRIOR APPLICATION NUMBER: 08/943,979
 ; PRIOR FILING DATE: October 4, 1997
 ; PRIOR APPLICATION NUMBER: 60/027,782
 ; PRIOR FILING DATE: October 4, 1996
 ; PRIOR APPLICATION NUMBER: 08/993,774
 ; PRIOR FILING DATE: December 18, 1997
 ; PRIOR APPLICATION NUMBER: 60/034,975
 ; PRIOR FILING DATE: December 20, 1996
 ; PRIOR APPLICATION NUMBER: 09/250,003
 ; PRIOR FILING DATE: February 10, 1999
 ; PRIOR APPLICATION NUMBER: 60/074,364
 ; PRIOR FILING DATE: February 12, 1998
 ; PRIOR APPLICATION NUMBER: 09/452,747
 ; PRIOR FILING DATE: December 1, 1999
 ; PRIOR APPLICATION NUMBER: 60/111,910
 ; PRIOR FILING DATE: December 10, 1998
 ; NUMBER OF SEQ ID NOS: 19698
 ; SOFTWARE: PERL Program
 ; SEQ ID NO: 8235
 ; LENGTH: 272
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No: hu00266046
 ; US-09-539-800C-8235

alignment_scores:

Quality: 7.00 Length: 7
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-09-539-800C-8235/rev

Align seg 1/1 to reverse of: US-09-539-800C-8235 from: 1 to: 272

7 ArgAlaSpSerArgPropio 13
 |||||
 122 AGGGCTGATCTCGGCCCCC 102

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-15360

seq_documentation_block:

; Sequence 15360, Application US/09540210B
 ; GENERAL INFORMATION:
 ; APPLICANT: Seilhamer, Jeffrey J.
 ; APPLICANT: Deleageane, Angelo M.
 ; APPLICANT: Stuart, Susan G.
 ; APPLICANT: Stuve, Laura L.
 ; APPLICANT: Mullahy, Sara J.
 ; APPLICANT: Naughton, Rebecca E.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
 ; FILE REFERENCE: PD-1037 CIP
 ; CURRENT APPLICATION NUMBER: US/09/540,210B
 ; CURRENT FILING DATE: 2002-04-03
 ; PRIOR APPLICATION NUMBER: 08/972,899
 ; PRIOR FILING DATE: November 18, 1997
 ; PRIOR APPLICATION NUMBER: 08/395,244
 ; PRIOR FILING DATE: February 27, 1995
 ; PRIOR APPLICATION NUMBER: 08/722,922

PRIOR FILING DATE: September 27, 1996
PRIOR APPLICATION NUMBER: 60/005,526
PRIOR FILING DATE: September 29, 1995
PRIOR APPLICATION NUMBER: 08/824,029
PRIOR FILING DATE: March 25, 1997
PRIOR APPLICATION NUMBER: 60/014,010
PRIOR FILING DATE: March 25, 1996
PRIOR APPLICATION NUMBER: 08/826,847
PRIOR FILING DATE: April 10, 1997
PRIOR APPLICATION NUMBER: 60/015,533
PRIOR FILING DATE: April 10, 1996
PRIOR APPLICATION NUMBER: 08/903,555
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/023,308
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/862,178
PRIOR FILING DATE: May 22, 1997
PRIOR APPLICATION NUMBER: 60/018,217
PRIOR FILING DATE: May 23, 1996
PRIOR APPLICATION NUMBER: 08/881,589
PRIOR FILING DATE: June 24, 1997
PRIOR APPLICATION NUMBER: 60/021,275
PRIOR FILING DATE: June 25, 1996
PRIOR APPLICATION NUMBER: 08/903,802
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/023,308
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/905,881
PRIOR FILING DATE: August 1, 1997
PRIOR APPLICATION NUMBER: 60/025,204
PRIOR FILING DATE: August 1, 1996
PRIOR APPLICATION NUMBER: 08/903,471
PRIOR FILING DATE: July 30, 1997
PRIOR APPLICATION NUMBER: 60/025,478
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/903,556
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/025,217
PRIOR FILING DATE: August 22, 1996
PRIOR APPLICATION NUMBER: 08/937,142
PRIOR FILING DATE: September 23, 1997
PRIOR APPLICATION NUMBER: 60/026,598
PRIOR FILING DATE: September 24, 1996
PRIOR APPLICATION NUMBER: 08/960,746
PRIOR FILING DATE: October 29, 1997
PRIOR APPLICATION NUMBER: 60/030,144
PRIOR FILING DATE: October 30, 1996
PRIOR APPLICATION NUMBER: 08/826,847
PRIOR FILING DATE: April 10, 1997
PRIOR APPLICATION NUMBER: 60/015,533
PRIOR FILING DATE: April 10, 1996
PRIOR APPLICATION NUMBER: 08/755,524
PRIOR FILING DATE: November 22, 1996
PRIOR APPLICATION NUMBER: 60/007,495
PRIOR FILING DATE: November 22, 1995
PRIOR APPLICATION NUMBER: 09/021,031
PRIOR FILING DATE: February 10, 1998
PRIOR APPLICATION NUMBER: 60/039,325
PRIOR FILING DATE: February 13, 1997
PRIOR APPLICATION NUMBER: 09/035,172
PRIOR FILING DATE: March 4, 1998
PRIOR APPLICATION NUMBER: 60/040,431
PRIOR FILING DATE: March 5, 1997
PRIOR APPLICATION NUMBER: 09/041,894
PRIOR FILING DATE: March 12, 1998
PRIOR APPLICATION NUMBER: 60/040,199
PRIOR FILING DATE: March 14, 1997
PRIOR APPLICATION NUMBER: 09/050,817
PRIOR FILING DATE: March 30, 1998
PRIOR APPLICATION NUMBER: 60/043,792
PRIOR FILING DATE: April 11, 1997
PRIOR APPLICATION NUMBER: 09/074,999
PRIOR FILING DATE: May 8, 1998

PRIOR APPLICATION NUMBER: 60/048,431
PRIOR FILING DATE: May 29, 1997
PRIOR APPLICATION NUMBER: 09/107,592
PRIOR FILING DATE: June 30, 1998
PRIOR APPLICATION NUMBER: 60/052,751
PRIOR FILING DATE: July 1, 1997
PRIOR APPLICATION NUMBER: 09/094,079
PRIOR FILING DATE: June 9, 1998
PRIOR APPLICATION NUMBER: 60/049,975
PRIOR FILING DATE: June 13, 1997
NUMBER OF SEQ ID NOS: 35654
SOFTWARE: PERL Program
SEQ ID NO 15360
LENGTH: 279
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: hu00456393
US-09-540-210B-15360

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-540-210B-15360/rev ..
Align seg 1/1 to reverse of: US-09-540-210B-15360 from: 1 to: 279

62 ThrsEulerSerLeuArgSer 68
.....
45 ACCAGCCTCTCCCTGAGGAGC 25

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-540-210B-20652

seq_documentation_block:
Sequence 20652, Application US/09540210B
GENERAL INFORMATION:
APPLICANT: Sellhammer, Jeffrey J.
APPLICANT: Deleogene, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Muliahy, Sara J.
TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
FILE REFERENCE: PD-1037 CIP
CURRENT APPLICATION NUMBER: US/09/540,210B
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: 08/972,899
PRIOR FILING DATE: November 18, 1997
PRIOR APPLICATION NUMBER: 08/395,244
PRIOR FILING DATE: February 27, 1995
PRIOR APPLICATION NUMBER: 08/722,922
PRIOR FILING DATE: September 27, 1996
PRIOR APPLICATION NUMBER: 60/005,526
PRIOR FILING DATE: September 29, 1995
PRIOR APPLICATION NUMBER: 08/824,029
PRIOR FILING DATE: March 25, 1997
PRIOR APPLICATION NUMBER: 60/014,010
PRIOR FILING DATE: March 25, 1996
PRIOR APPLICATION NUMBER: 08/826,847
PRIOR FILING DATE: April 10, 1997
PRIOR APPLICATION NUMBER: 60/015,533
PRIOR FILING DATE: April 10, 1996
PRIOR APPLICATION NUMBER: 08/903,555
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/023,308
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/862,178
PRIOR FILING DATE: May 22, 1997


```

; PRIOR APPLICATION NUMBER: 60/018, 217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 20652
; LENGTH: 282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: hu00094015
; NAME/KEY: unsure
; LOCATION: 189-190, 229
; OTHER INFORMATION: a, t, c, g, or other
; US-09-540-210B-20652

```

```

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
  US-09-528-682-1 x US-09-540-210B-20652/rev ..

```

```

Align seg 1/1 to reverse of: US-09-540-210B-20652 from: 1 to: 282

```

```

66 leuArgSerAlaHisLeuAla 72
|||||
75 TTACGGCTCTGCACACCTGGCC 55

```

```

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-28303

```

```

seq_documentation_block:
; Sequence 28303, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara L.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; PRIOR FILING DATE: 2000-03-30
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 28303
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: hu00114685
; US-09-539-331D-28303

```

```

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

```

alignment_block:
  US-09-528-682-1 x US-09-539-331D-28303 ..

```

```

Align seg 1/1 to: US-09-539-331D-28303 from: 1 to: 312

```

```

63 SerLeuSerLeuArgSerAla 69
|||||
31 AGCCTGAGTCTGAGAGTGC 51

```

```

seq_name: /cgn2.6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-331D-22417

```

```

seq_documentation_block:
; Sequence 22417, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.

```



```
; APPLICANT: Mullaby, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539, 331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 22417
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID NO: hu00325793
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 45, 51, 64, 78, 93, 108, 144, 150, 168, 216, 234, 249
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-331D-22417

alignment_scores:
  Quality: 7.00      length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-539-331D-22417/rev ..

Align seg 1/1 to reverse of: US-09-539-331D-22417 from: 1 to: 318

66 leuArgSerLeuAlaHisLeuAla 72
|||||
206 CTAAAGCTCGCGCAGCTTGCA 186

seq_name: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:PCT-US02-10421-160

seq_documentation_block:
; Sequence 160, Application PC/TUS0210421
; GENERAL INFORMATION:
; APPLICANT: Corixa Corporation
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Siqing
; APPLICANT: Chaitanya S.
; APPLICANT: Gaiger, Alexander
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF CANCER
; FILE REFERENCE: 210121.565PC
; CURRENT APPLICATION NUMBER: PCT/US02/10421
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2959
; SEQ ID NO 160
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-10421-160

alignment_scores:
  Quality: 7.00      length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x PCT-US02-10421-160 ..

Align seg 1/1 to: PCT-US02-10421-160 from: 1 to: 323

63 SerLeuSerLeuArgSerAla 69
|||||
71 TCTCTTCTCTCTCAGAAAGTGCT 91
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-15587

seq_documentation_block:
; Sequence 15587, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Afterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jessen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroja, Mimi
; APPLICANT: Lomeili, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raist, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fel
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/721, 544
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515, 128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034, 341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15587
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-544-15587

alignment_scores:
  Quality: 7.00      length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-721-544-15587/rev ..

Align seg 1/1 to reverse of: US-09-721-544-15587 from: 1 to: 323

61 SerThrSerLeuSerLeuArg 67
|||||
263 TCAACATCACTGCTCTTGAGG 243

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-112-699-160

seq_documentation_block:
; Sequence 160, Application US/10112699
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Siqing
; APPLICANT: Chaitanya S.
```



```
; APPLICANT: Gaiger, Alexander
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.565
; CURRENT APPLICATION NUMBER: US/10/112.699
; NUMBER OF SEQ ID NOS: 2959
; SEQ ID NO 160
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-112-699-160

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-112-699-160 ..

Align seg 1/1 to: US-10-112-699-160 from: 1 to: 323

63 SerLeuSerLeuArgSerAla 69
|||||
71 TCTCTTCTCTCAGAGTGCT 91

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800C-15776

seq_documentation_block:
; Sequence 15776, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539.800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521.383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271.217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334.881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943.978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028.732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943.979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027.782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993.774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034.975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250.003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074.364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452.747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111.910
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 15776
; LENGTH: 324
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu01286321
; US-09-539-800C-15776

alignment_scores:
    Quality: 7.00      Length: 7
    Ratio: 1.000      Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-539-800C-15776/rev ..

Align seg 1/1 to reverse of: US-09-539-800C-15776 from: 1 to: 324

112 GluValSerAlaLeuGly 118
|||||
85 GAGTAAGTGCACCTTGAGGG 65

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-539-800B-4227

seq_documentation_block:
; Sequence 4227, Application US/09539806B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE
; FILE REFERENCE: PD-1027 CIP
; CURRENT APPLICATION NUMBER: US/09/539.806B
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/706.766
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/004.676
; PRIOR FILING DATE: October 2, 1995
; PRIOR APPLICATION NUMBER: 08/749.515
; PRIOR FILING DATE: November 15, 1996
; PRIOR APPLICATION NUMBER: 60/006.810
; PRIOR FILING DATE: November 15, 1995
; PRIOR APPLICATION NUMBER: 08/822.285
; PRIOR FILING DATE: March 20, 1997
; PRIOR APPLICATION NUMBER: 60/013.696
; PRIOR FILING DATE: March 20, 1996
; PRIOR APPLICATION NUMBER: 08/951.197
; PRIOR FILING DATE: October 1, 1997
; PRIOR APPLICATION NUMBER: 60/027.249
; PRIOR FILING DATE: October 1, 1996
; PRIOR APPLICATION NUMBER: 08/826.438
; PRIOR FILING DATE: March 20, 1997
; PRIOR APPLICATION NUMBER: 60/016.145
; PRIOR FILING DATE: April 18, 1996
; PRIOR APPLICATION NUMBER: 60/013.696
; PRIOR FILING DATE: March 20, 1996
; PRIOR APPLICATION NUMBER: 08/839.389
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 60/015.312
; PRIOR FILING DATE: April 12, 1996
; PRIOR APPLICATION NUMBER: 08/951.198
; PRIOR FILING DATE: October 1, 1997
; PRIOR APPLICATION NUMBER: 60/027.249
; PRIOR FILING DATE: October 1, 1996
; PRIOR APPLICATION NUMBER: 08/839.968
; PRIOR FILING DATE: April 24, 1997
; PRIOR APPLICATION NUMBER: 60/016.150
; PRIOR FILING DATE: April 24, 1996
; PRIOR APPLICATION NUMBER: 08/846.104
; PRIOR FILING DATE: April 25, 1997
```



```

? PRIOR APPLICATION NUMBER: 60/016,838
? PRIOR FILING DATE: April 26, 1996
? PRIOR APPLICATION NUMBER: 08/999,886
? PRIOR FILING DATE: May 13, 1997
? PRIOR APPLICATION NUMBER: 60/017,555
? PRIOR FILING DATE: May 13, 1996
? PRIOR APPLICATION NUMBER: 60/031,667
? PRIOR FILING DATE: November 11, 1996
? PRIOR APPLICATION NUMBER: 08/901,902
? PRIOR FILING DATE: July 24, 1997
? PRIOR APPLICATION NUMBER: 60/023,379
? PRIOR FILING DATE: July 26, 1996
? PRIOR APPLICATION NUMBER: 08/918,181
? PRIOR FILING DATE: August 27, 1997
? PRIOR APPLICATION NUMBER: 60/023,236
? PRIOR FILING DATE: August 28, 1996
? PRIOR APPLICATION NUMBER: 08/940,864
? PRIOR FILING DATE: September 29, 1997
? PRIOR APPLICATION NUMBER: 60/027,236
? PRIOR FILING DATE: September 30, 1996
? PRIOR APPLICATION NUMBER: 08/956,502
? PRIOR FILING DATE: October 22, 1997
? PRIOR APPLICATION NUMBER: 60/029,083
? PRIOR FILING DATE: October 23, 1996
? PRIOR APPLICATION NUMBER: 08/993,402
? PRIOR FILING DATE: December 18, 1997
? PRIOR APPLICATION NUMBER: 60/033,647
? PRIOR FILING DATE: December 19, 1996
? PRIOR APPLICATION NUMBER: 09/022,355
? PRIOR FILING DATE: February 11, 1998
? PRIOR APPLICATION NUMBER: 60/044,847
? PRIOR FILING DATE: February 13, 1997
? PRIOR APPLICATION NUMBER: 09/045,574
? PRIOR FILING DATE: March 20, 1998
? PRIOR APPLICATION NUMBER: 60/041,275
? PRIOR FILING DATE: March 21, 1997
? PRIOR APPLICATION NUMBER: 09/042,629
? PRIOR FILING DATE: March 16, 1998
? PRIOR APPLICATION NUMBER: 60/043,613
? PRIOR FILING DATE: March 18, 1997
? PRIOR APPLICATION NUMBER: 09/057,988
? PRIOR FILING DATE: April 9, 1998
? PRIOR APPLICATION NUMBER: 60/043,256
? PRIOR FILING DATE: April 16, 1997
? PRIOR APPLICATION NUMBER: 09/114,041
? PRIOR FILING DATE: July 10, 1998
? PRIOR APPLICATION NUMBER: 60/052,257
? PRIOR FILING DATE: July 10, 1997
? PRIOR APPLICATION NUMBER: 09/145,340
? PRIOR FILING DATE: September 1, 1998
? PRIOR APPLICATION NUMBER: 60/058,922
? PRIOR FILING DATE: September 11, 1997
? NUMBER OF SEQ ID NOS: 48372
? SOFTWARE: PERL Program
? SEQ ID NO 4227
? LENGTH: 349
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: Incyte ID No: hu00577412
? FEATURE:
? NAME/KEY: unsure
? LOCATION: 37, 56, 60, 62, 66, 69
? OTHER INFORMATION: a, t, c, g, or other
US-09-539-806B-4227

```

```

alignment_scores:
  Quality: 7.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 7
  Gaps: 0
  Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x US-09-539-806B-4227 ..
Align seg 1/1 to: US-09-539-806B-4227 from: 1 to: 349

```

```

114 SerIaLeuGIgLyIlePro 120
|||||
264 TCTGCATTAGCGGCATCCCA 284

```

```
seq_name: /cgn2_6/prodata/1/pna/us09_NEW_COMB.seq:US-09-721-544-19743

```

```
seq_documentation_block:

```

```
; Sequence 19743, Application US/09721544
; GENERAL INFORMATION:

```

```

; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jensen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroja, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Ralsi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fel
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/721,544
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515,128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034,341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19743
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-544-19743

```

```

alignment_scores:
  Quality: 7.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 7
  Gaps: 0
  Percent Identity: 100.000

```

```

alignment_block:
US-09-528-682-1 x US-09-721-544-19743 ..

```

```
Align seg 1/1 to: US-09-721-544-19743 from: 1 to: 356

```

```

61 SerThrSerLeuSerLeuArg 67
|||||
27 AGCACTTCATCAGTTAAGA 47

```


seq_name: /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-011-154-2094

seq_documentation_block:

```
; Sequence 2094, Application US/10011154
; GENERAL INFORMATION:
; APPLICANT: Dmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark C.
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 780CIP
; CURRENT APPLICATION NUMBER: US/10/011,154
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/524,038
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/404,284
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 4670
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 2094
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-011-154-2094
```

alignment_scores:

Quality:	7.00	Length:
Ratio:	1.000	Gaps: 0
Percent Similarity:	100.000	Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-011-154-2094/rev ..

Align seg 1/1 to reverse of: US-10-011-154-2094 from: 1 to: 410

```
65 SerleuArgSerAlaHisLeu 71
|||||
353 TCCCTTGCTGCTGTCATCTC 333
```

seq_name: /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-3305

seq_documentation_block:

```
; Sequence 3305, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3305
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Human
```

US-10-027-632-3305

alignment_scores:

Quality:	7.00	Length:
Ratio:	1.000	Gaps: 0
Percent Similarity:	100.000	Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-027-632-3305/rev ..

Align seg 1/1 to reverse of: US-10-027-632-3305 from: 1 to: 435

```
61 SerThrSerLeuSerLeuArg 67
|||||
103 TCAACTGCTTGAGCCGACCA 83
```

seq_name: /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-011-154-2817

seq_documentation_block:

```
; Sequence 2817, Application US/10011154
; GENERAL INFORMATION:
; APPLICANT: Dmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark C.
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 780CIP
; CURRENT APPLICATION NUMBER: US/10/011,154
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/524,038
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/404,284
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 4670
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 2817
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-011-154-2817
```

alignment_scores:

Quality:	7.00	Length:
Ratio:	1.000	Gaps: 0
Percent Similarity:	100.000	Percent Identity: 100.000

alignment_block:

US-09-528-682-1 x US-10-011-154-2817 ..

Align seg 1/1 to: US-10-011-154-2817 from: 1 to: 436

```
65 SerLeuArgSerAlaHisLeu 71
|||||
283 TCCCTTGCTGCTGTCATCTC 303
```

seq_name: /cgn2.6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-89412

seq_documentation_block:

```
; Sequence 89412, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```



```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89412
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-89412
```

```
alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-1 x US-10-027-632-89412  ..
```

```
Align seg 1/1 to: US-10-027-632-89412 from: 1 to: 448
```

```
60 ValSerThrSerLeuSerLeu 66
|||||
335 GTGTCTACATCTCTCTCTC 355
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-89413
```

```
seq_documentation_block:
; Sequence 89413, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89413
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-89413
```

```
alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-1 x US-10-027-632-89413  ..
```

```
Align seg 1/1 to: US-10-027-632-89413 from: 1 to: 448
```

```
60 ValSerThrSerLeuSerLeu 66
|||||
335 GTGTCTACATCTCTCTCTC 355
```

```
seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-144-702-194
```

```
seq_documentation_block:
; Sequence 194, Application US/10144702
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PMO10CIN
; CURRENT APPLICATION NUMBER: US/10/144,702
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/758,460
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 940
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 194
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (395)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (456)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (457)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-144-702-194
```

```
alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
US-09-528-682-1 x US-10-144-702-194/rev  ..
```

```
Align seg 1/1 to reverse of: US-10-144-702-194 from: 1 to: 457
```

```
62 ThrSerLeuArgSer 68
|||||
99 ACTTCCCTGAGCCTCAGTCC 79
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-2302
```

```
seq_documentation_block:
; Sequence 2302, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simln
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
```



```

1  APPLICANT: Garcia, Veronica
2  APPLICANT: Giedt, Gretchen
3  APPLICANT: Hunter, Kelly
4  APPLICANT: Jessen, Aaron
5  APPLICANT: Jones, Lee
6  APPLICANT: Kita, David
7  APPLICANT: Labat, Ivan
8  APPLICANT: Laroya, Mimi
9  APPLICANT: Lomelli, Michelle
10 APPLICANT: Nguyen, Phuong
11 APPLICANT: Nogra, Margie
12 APPLICANT: Palencia, Servando
13 APPLICANT: Raissi, Fariba
14 APPLICANT: Smith, Benjamin
15 APPLICANT: Tkach, Joe
16 APPLICANT: Tran, Lien
17 APPLICANT: Verna, Ron
18 APPLICANT: Yang, Fel
19 APPLICANT: Yim, Kenneth
20 TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
21 FROM A CDNA LIBRARY OF FETAL LIVER-SPLEEN
22 FILE REFERENCE: 728CIP
23 CURRENT APPLICATION NUMBER: US/09/721,544
24 CURRENT FILING DATE: 2000-11-21
25 PRIOR APPLICATION NUMBER: 09/515,128
26 PRIOR FILING DATE: 1998-02-13
27 PRIOR APPLICATION NUMBER: 09/034,341
28 PRIOR FILING DATE: 1998-02-13
29 NUMBER OF SEQ ID NOS: 24489
30 SOFTWARE: FastSeq for Windows Version 3.0
31 SEQ ID NO 2302
32 LENGTH: 470
33 TYPE: DNA
34 ORGANISM: Homo sapiens
35 US-09-721-544-2302

```

```
alignment_scores:      7
                    quality: 7
                    Ratio: 1.000
                    Gaps: 0
Percent Similarity: 100.000
Percent Identity: 100.000
```

alignment_block: NS-09-528-682-1 x NS-09-721-544-2302

Align seg 1/1 to: US-09-721-544-2302 from: 1 to: 470

187 CysGlyAsnSerSerArgThr 193
 11111111111111111111
 259 TGTGGCAATTCCTCAAGACC 279

seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-704-302A-158

```

seq documentation block:
; Sequence 158. Application US/09704302A
;
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c10
; CURRENT APPLICATION NUMBER: US/09/704, 302A
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 1402
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 158
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-09-704-302A-158

```

alignment_scores:

Quality:	7.00	Length:	7
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block: MS-09-528-682-1 x MS-09-704-302A-158/rev

Align seg 1/1 to reverse of: US-09-704-302A-158 from: 1 to: 471

```

71 LeuAlaGlyInserIIeLeu 77
   |||||
361 CTTGCCGGACAATCCATTCTC 341

```

seq_name: /cyn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-002-623-355

```
seq_documentation_block:
  Sequence 355, Application US/10002623
  GENERAL INFORMATION:
    APPLICANT: OEFNER, PETER J.
    APPLICANT: UNDERHILL, PETER A.
    TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
    TITLE OF INVENTION: AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
    TITLE OF INVENTION: POPULATIONS
    FILE REFERENCE: STAN-212
    CURRENT APPLICATION NUMBER: US/10/002,623
    CURRENT FILING DATE: 2001-11-01
    PRIOR APPLICATION NUMBER: US 60/245,355
    PRIOR FILING DATE: 2000-11-01
    NUMBER OF SEQ ID NOS: 952
    SOFTWARE: FastSeq for Windows Version 4.0
    SEQ ID NO 355
    LENGTH: 478
    TYPE: DNA
    ORGANISM: Homo Sapiens
  US-10-002-623-355
```

```
alignment_scores:
  Quality: 7.00
  Ratio: 1.000
  Gaps: 0
Percent Similarity: 100.000
Percent Identity: 100.000
```

```
alignment_block:
  mc-10-003-623-355 crew
  mc-09-538-683-1
```

Align seg 1/1 to reverse of: US-10-002-623-355 from: 1 to: 478

	GluGlnValSerAlaLeu	116
110		
	GAGCAGGAGGTCTCAGCTTA	303
323		

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-86877

```

seq documentation block:
? Sequence 86877, Application US/10027632
? GENERAL INFORMATION:
? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single Nucleotides
? TITLE OF INVENTION: Polymorphisms in the Human Genome
? FILE REFERENCE: 108827.129
? CURRENT APPLICATION NUMBER: US/10/027,632
? CURRENT FILING DATE: 2002-04-30
? PRIOR APPLICATION NUMBER: US 60/218,006
? PRIOR FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 60/198,676
? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193,483
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185,218
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/155,358

```

alignment_scores:


```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86877
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-86877

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-86877  ..

Align seg 1/1 to: US-10-027-632-86877 from: 1 to: 478
144 GUTYTAAGAspArgTyrTyr 150
|||||
224 GAATATAGAGACAGTATTAT 244

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-86878

seq_documentation_block:
; Sequence 86878, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: US/10/027,632
; PRIOR APPLICATION NUMBER: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86878
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-86878

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-86878  ..

Align seg 1/1 to: US-10-027-632-86878 from: 1 to: 478
144 GUTYTAAGAspArgTyrTyr 150
|||||
224 GAATATAGAGACAGTATTAT 244
```

```
seq_name: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:US-09-721-544-3385

seq_documentation_block:
; Sequence 3385, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vada
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radcoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Gledt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jensen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroja, Mimi
; APPLICANT: Lomeili, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisl, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Rei
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 728CIP
; CURRENT FILING DATE: From a CDNA Library of Fetal Liver-Spleen
; CURRENT APPLICATION NUMBER: US/09/721,544
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515,128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034,341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3385
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(488)
; OTHER INFORMATION: n = A,T,C or G
US-09-721-544-3385

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-09-721-544-3385/rev  ..

Align seg 1/1 to reverse of: US-09-721-544-3385 from: 1 to: 488
62 ThrSerLeuSerLeuArgSer 68
|||||
265 ACAATCAATTAGCCTTCGAACT 245

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-68521

seq_documentation_block:
; Sequence 68521, Application US/10027632
```



```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68521
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68521

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-68521/rev ..

Align seg 1/1 to reverse of: US-10-027-632-68521 from: 1 to: 493

208 ThrTleTyLeuArgGluTyr 214
|||||
403 ACTATCTATTTAAGAGAGTAC 383

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-295156
seq_documentation_block:
; Sequence 295156, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 295156
; LENGTH: 493
```

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-295156

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-295156/rev ..

Align seg 1/1 to reverse of: US-10-027-632-295156 from: 1 to: 493

208 ThrTleTyLeuArgGluTyr 214
|||||
403 ACTATCTATTTAAGAGAGTAC 383

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-60190
seq_documentation_block:
; Sequence 60190, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60190
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-60190

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-528-682-1 x US-10-027-632-60190/rev ..

Align seg 1/1 to reverse of: US-10-027-632-60190 from: 1 to: 502

58 GLTYrValSerThrSerLeu 64
|||||
485 GGGTATGTTTCTACTTCACGC 465

seq_name: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:US-10-027-632-298567
seq_documentation_block:
; Sequence 298567, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```



```

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 298567
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-298567

```

```

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment_block:

US-09-528-682-1 x US-10-027-632-298567/rev ..

Align seg 1/1 to reverse of: US-10-027-632-298567 from: 1 to: 502

```

58 GlyTyrValSerThrSerLeu 64
|||||
485 GGGTATGTTCTACTCAGTC 465

```


